

AN APPLICATION OF MULTIVARIATE STATISTICAL TECHNIQUES TO THE ANALYSIS
OF THE OPERATIONAL EFFECTIVENESS OF A MILITARY FORCE

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

James Thomas Baird

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Georgia Institute of Technology

May, 1976

AN APPLICATION OF MULTIVARIATE STATISTICAL TECHNIQUES TO THE ANALYSIS
OF THE OPERATIONAL EFFECTIVENESS OF A MILITARY FORCE

Approved:

Chairman, Russell G. Heikes

Harrison M. Wadsworth

Douglas C. Montgomery

Richard D. Wright

Date approved by Chairman: 5-26-76

ACKNOWLEDGMENTS

I would like to express my appreciation to the United States Army for providing me the opportunity to attend graduate school and to the Army Materiel Systems Analysis Agency for extending contractual support for this research.

My sincerest appreciation goes also to Dr. Russell G. Heikes, the chairman of my thesis committee, to Dr. Harry M. Wadsworth, Dr. Douglas C. Montgomery, and to Mr. Richard D. Wright, the members of my committee, for their assistance and guidance throughout this research effort. To Mr. Frank B. Alt, I extend a special expression of gratitude for explaining the rudiments of multivariate statistics.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
SUMMARY.	vii
Chapter	
I. INTRODUCTION	1
Background	
The Army System Acquisition Process	
Operational Testing	
Nature and Scope of the Problem	
Mobility	
Objective and Procedure	
II. DEFINITIONS OF OPERATIONAL MOBILITY.	9
Introduction	
Current Concepts	
Macro-view of Operational Mobility	
Micro-view of Operational Mobility	
Proposed Definitions	
Measurement of Relative Momentum	
Relative Momentum as a MOE to Describe Operational Mobility	
Structure of the Problem	
III. REVIEW OF APPLICABLE MULTIVARIATE STATISTICAL METHODS. .	19
Introduction	
Basic Multivariate Statistical Concepts and Notation	
The Assumption of Multivariate Normality	
Tests for Multivariate Normality	
Data Transformation	
Tests of Hypotheses	
Tests of Hypotheses on Mean Vectors	
Simultaneous Confidence Intervals	
Statistical Inference After Non-Linear Transformation	
Principal Components	
Contribution of Individual Variables	
Multivariate Random Vector Generation	

	Page
IV. DEVELOPMENT OF THE SPECIFIC METHODOLOGY	57
Introduction	
Determination of Significant Differences	
Contribution of Individual MOE	
Methodological Input	
V. DEMONSTRATION OF THE METHODOLOGY	66
Introduction	
Background of the Problems	
Example Problem I	
Section I	
Section II	
Example Problem II	
Further Analysis of MOE Contributions	
VI. SUMMARY AND RECOMMENDATIONS	101
APPENDIX A	104
APPENDIX B	107
BIBLIOGRAPHY	139

LIST OF TABLES

Table		Page
1.	Order Statistics for $X_{\sim 1}$ and $X_{\sim 2}$	69
2.	Confidence Intervals for $(\mu_{Y1j} - \mu_{0j}^T)$	73
3.	Confidence Intervals for $(\mu_{X1j} - \mu_0)$	74
4.	Confidence Intervals for $(\mu_{Y1j} - \mu_{Y2j})$	74
5.	Principal Components for $X_{\sim 1}$	81
6.	MOE Contributions for CFV-1 (%)	82
7.	Principal Components for $X_{\sim 2}$	82
8.	MOE Contributions for CFV-2 (%)	82
9.	Principal Components for $(X_{\sim 1} - X_{\sim 2})$	83
10.	MOE Contributions for Differences (%)	83
11.	Confidence Intervals for $(\mu_{X1j} - \mu_{0j})$	85
12.	Confidence Intervals for $(\mu_{Y1j} - \mu_{Y2j})$	86
13.	Operational Mobility Contributions (%)	95
14.	Principal Components for $X_{\sim 1}$	96
15.	Principal Components for $X_{\sim 2}$	97
16.	Principal Components for $(X_{\sim 1} - X_{\sim 2})$	98
17.	MOE Contributions for CFV-1 (%)	99
18.	MOE Contributions for CFV-2 (%)	99
19.	MOE Contributions for Differences (%)	99
20.	Best Linear Coefficients, $\{a_{n-1+1}\}$, for $n = 11, 12, \dots, 20$	105
21.	Percentage Points of W for $n = 3, 4, \dots, 20$	106

LIST OF ILLUSTRATIONS

Figure	Page
1a. Q Region for $(\mu_{X1j} - \mu_{X2j})$	41
1b. Establishing Intervals for H_0 True	42
1c. G Region for $(\mu_{X1j} - \mu_{X2j})$	43
1d. Q and G Regions for $(\mu_{X1j} - \mu_{X2j})$	44
2. Scatter Diagram of Bivariate Observations	49
3. Flow Diagram of the Methodology	64
4. Q and G Regions for $(\mu_{X11} - \mu_{X21})$	76
5. Q and G Regions for $(\mu_{X12} - \mu_{X22})$	77
6. Q and G Regions for $(\mu_{X13} - \mu_{X23})$	78
7. Q and G Regions for $(\mu_{X14} - \mu_{X24})$	79
8. Q and G Regions for $(\mu_{X11} - \mu_{X21})$	87
9. Q and G Regions for $(\mu_{X12} - \mu_{X22})$	88
10. Q and G Regions for $(\mu_{X13} - \mu_{X23})$	89
11. Q and G Regions for $(\mu_{X14} - \mu_{X24})$	90
12. Q and G Regions for $(\mu_{X15} - \mu_{X25})$	91
13. Q and G Regions for $(\mu_{X16} - \mu_{X26})$	92
14. Q and G Regions for $(\mu_{X17} - \mu_{X27})$	93
15. Q and G Regions for $(\mu_{X18} - \mu_{X28})$	94

SUMMARY

This research addresses the problem of determining the contribution of operational mobility to the operational effectiveness of a military force and the contribution of differences in operational mobility to differences in operational effectiveness for forces employing competing land combat vehicle systems. Two definitions of the operational mobility of a force are proposed, along with suggested methods for their representation as quantifiable measures of effectiveness (MOE). The first definition is based on a measure of relative momentum between two forces engaged in combat, while the second is based on relative measures of critical operational performance characteristics of particular vehicles employed in combatant forces.

A general methodology, based on classical, well-documented, multivariate statistical procedures, is then developed that enables determination of significant differences between the operational effectiveness of forces employing competing materiel systems and relative contributions of individual MOE, to include any measure of operational mobility, to a force's operational effectiveness and differences between MOE to differences in operational effectiveness.

The methodology includes an investigation of the multivariate normality of a set of MOE obtained from hypothesized replications of a stochastic combat model by utilizing two previously developed tests for multivariate normality. One test is based on order statistics, while the other utilizes a non-linear transformation procedure that induces multi-

variate normality into a data set. An approximate procedure was also developed to facilitate translation of inferences drawn after non-linear transformation to inferences on original variables.

The methodology is applied to two hypothetical example problems to illustrate the fact that either definition of operational mobility is conformable to analysis. Statistical analyses in each example problem are limited to the two sample case with equal sample sizes, unknown covariance matrices, and assumed independence between data sets.

CHAPTER I

INTRODUCTION

Background

The Army System Acquisition Process

The development, testing, and final acquisition of major materiel systems by the Department of the Army is an evolutionary process designed to insure that the United States Army possesses the best equipment available in order to facilitate the accomplishment of its mission. Specific guidelines for implementation of this procedure are contained in appropriate Army Regulations that follow basic policies established by the Department of Defense (41). A detailed description of the Army materiel acquisition procedure will not be presented in this study; however, a brief description will be presented so that the framework of the problem examined herein may be established. The interested reader is referred to Bettencourt (9) and Burnette (13) for a detailed summary of the process from the conceptual development phase through validation, full scale development, and final production and deployment of the particular system.

Each of the first three phases mentioned above consists of independently conducted operational and developmental tests. Operational Testing (OT) is concerned primarily with evaluating the operational effectiveness of a force in which the new system is utilized, while Developmental Testing (DT) examines the engineering design characteristics of the system itself. The results of these tests are forwarded first to the Army Systems Acquisition Review Council (ASARC), then to the Defense

Systems Acquisition Review Council (DSARC) at the conclusion of each phase of the acquisition process. The final decision with respect to proceeding to the next phase of testing, modifying the system and returning to the beginning of the previous phase, or terminating the project rests with the Secretary of Defense (44).

Operational Testing

This research is concerned primarily with the OT segment of each phase of the acquisition process. As mentioned previously, a primary objective of OT is the evaluation of the operational effectiveness of a military force in which a new system is employed. For example, if the Army had validated a requirement for a new, or improved, armored personnel carrier (APC), then OT would most probably examine the operational effectiveness of a battalion level mechanized infantry force.

Within the structure of Operational Testing, several critical issues arise. A suitable definition of operational effectiveness must first be established so that the military worth of the new system may be properly evaluated. Secondly, this definition must be able to be expressed in measurable quantities so that an adequate representation of operational effectiveness may be obtained from test results. Additionally, experiments must be designed so that the measurable quantities, usually termed Measures of Effectiveness (MOE), may be calculated from experimental data. Finally, appropriate statistical techniques must be utilized in order to evaluate a force's operational effectiveness in terms of the designated MOE so that viable and realistic comparisons may be made between the new system and the existing system. The results of these comparisons, together with subjective military evaluations, are forwarded

to ASARC and DSARC to supplement the appropriate decision maker's portfolio of information concerning the new system.

With respect to these critical issues, the Army has developed policies and procedures to insure that evaluation of a force's operational effectiveness is adequately conducted. Succinctly stated, the definition of operational effectiveness is the ability of a force to accomplish its assigned mission (48). Operational Testing, then, must evaluate a force's ability with a new system and compare this to the force's ability with the current system.

Based on this definition and on the nature of the systems and force to be examined, the appropriate test agency selects pertinent MOE that most adequately reflect the ability of the particular force to accomplish its mission. Specific MOE that may be examined will be discussed in a subsequent section.

After final approval of the MOE by the Commander, U. S. Army Operational Test and Evaluation Agency (OTEA), the experimental environment must be established. Traditionally, these experiments encompass field testing of active Army units in controlled maneuvers under simulated combat conditions. In addition to field tests, implementation of other test procedures is highly desirable in order to maximize the amount of information available to the decision maker.

An experimental procedure that is available to provide this supplemental information is a Cost and Operational Effectiveness Analysis (COEA). This type of analysis is a periodic review of the effectiveness of existing materiel systems traditionally based on analysis of deterministic and/or stochastic computer simulations of combat against a hy-

pothetical threat force (48). Recent advancements in computer technology, coupled with near prohibitive costs of producing large quantities of a major piece of equipment to be used solely for test purposes, make the COEA a highly desirable method of analysis.

Conducting a COEA concurrently with, and independently of, one of the phases of OT in the development of a new system of equipment can thus prove extremely advantageous in efforts to compare new and current materiel systems. In light of this, the research presented herein will be concerned with the analysis of operational effectiveness data obtained from stochastic computer simulations of hypothesized combat.

Once the appropriate test environment has been selected, tests are conducted to evaluate a force's operational effectiveness in varying battle conditions. An advantage of computer simulation in a COEA becomes evident here due to the fact that several battle scenarios may be examined (*e. g.*, attack, defense, or delay) in a variety of terrain settings (*e. g.*, desert or densely vegetated areas). Results of the simulation are then analyzed using appropriate statistical techniques so that viable comparative statements concerning a force's operational effectiveness with competing systems may be made. Historically, univariate techniques, such as the Analysis of Variance (ANOVA), are used for this purpose (47).

Nature and Scope of the Problem

The statistical analysis of a force's operational effectiveness must be able to detect significant differences between effectiveness with a current system and effectiveness with a proposed, or alternative, system. Similarly, information regarding differences in force effectiveness

between several proposed alternatives may also be desired. Analyses must also detect significant differences between corresponding MOE being examined so that differences in effectiveness may be further isolated and attributed to their proper source.

This research, therefore, will be directed toward the development of a methodology that will encompass the detection of differences in operational effectiveness from one system to another and the determination of the contribution of the MOE being examined to the existing differences. Although the methodology will be applicable to the analysis of a variety of military systems, this research will be concerned with the analysis of the operational effectiveness of a force in which the major system being examined is a land combat vehicle. Inherent in the examination of such a system is an assessment of the system's mobility and a determination of the contribution of mobility to the force's operational effectiveness.

Mobility

In examining the mobility of a land combat system more closely, it becomes apparent that there are in fact two types of mobility that must be assessed: performance mobility and operational mobility (42). Performance mobility may be described in terms of the design characteristics of the particular vehicle being examined. An assessment of the performance mobility of the vehicles of two competing systems then becomes merely a straightforward comparison of such characteristics as speed, range, acceleration, and ability to traverse obstacles. It should be noted, however, that a vehicle with superior performance characteristics may not in fact be the superior vehicle in terms of a force's oper-

ational effectiveness.

In light of this, the operational mobility of the vehicle must also be described and examined. Unfortunately, no precise definition of operational mobility exists. Heuristically, however, it may be viewed as the contribution of a system's mobility performance characteristics to the operational effectiveness of the force in which the system is employed. This contribution, when considered with other MOE in the description of operational effectiveness, would more adequately portray the operational effectiveness of a force that utilizes the particular land combat system. A single, or perhaps composite, measure of operational mobility is thus desired for a more comprehensive and rigorous analysis of systems of this type.

If it were possible to formulate a definition of operational mobility in terms of quantifiable MOE, then a statistical analysis of operational effectiveness might proceed as usual by utilizing univariate techniques. However, it has been recognized that correlations exist between MOE currently being used to describe operational effectiveness (42). For example, if two selected MOE were personnel losses and vehicle losses, then we might suspect the existence of a possible positive correlation. Although these MOE may in fact be independent, the potential presence of a correlated relationship must be explored. In a similar manner, a MOE describing operational mobility may also be related to personnel and vehicle losses. These possible relationships must also be exploited, which suggests utilizing multivariate statistical techniques, rather than univariate techniques, in an analysis of operational effectiveness.

Objective and Procedure

The primary objective of this research may now be stated in the following manner:

To develop an improved methodology utilizing multivariate statistical techniques to determine:

1. significant differences between the operational effectiveness of a force utilizing a current system of equipment and the operational effectiveness of a force utilizing a proposed, or alternative, system, and
2. relative contributions of individual MOE to operational effectiveness, and differences between MOE to differences in operational effectiveness.

Accomplishment of the stated objective for the land combat vehicle systems will necessarily include the determination not only of the contribution of operational mobility to the operational effectiveness of a military force, but also the contribution of differences in operational mobility to differences in operational effectiveness. These latter presumptions are highly contingent, of course, on the precise formulation of a realistic definition of operational mobility and qualification of this definition as a suitable MOE. The importance of this critical issue precipitates the necessity for its discussion immediately following a brief outline of the procedure to be followed throughout the remainder of this study.

The aforementioned procedure will consist of formulating two definitions of operational mobility, reviewing appropriate multivariate statistical techniques, developing the methodology, exhibiting the meth-

odology in two example problems, and discussing areas in which the methodology may be applied. Pertinent assumptions and limiting factors concerning the methodology and its application will be discussed separately and summarized in the concluding chapter.

CHAPTER II

DEFINITIONS OF OPERATIONAL MOBILITY

Introduction

Based on the results of previously conducted independent studies relating to concepts of operational mobility as a force attribute, two definitions of a force's operational mobility will be presented in this chapter. These definitions represent a synthesis of ideas, the primary objective in their formulation being the development of realistic, viable definitions suitable for quantification as one, or several, MOE. Expression of operational mobility in such a manner will obviously facilitate the calculation of its contribution to the operational effectiveness of a force.

Current Concepts

Macro-view of Operational Mobility

An excellent historical conceptualization of operational mobility was presented by the Research Analysis Corporation in 1967 (2). Although this study did not directly propose specific methods for quantification of the attribute in question, several practical concepts were introduced and subjectively examined in numerous battles from the era of Ghenghis Khan through the Korean War. Significant results of this study concerning a conceptual description of operational mobility may be summarized as follows:

1. Although individual and small unit mobility contribute to bat-

the outcome, it is the mobility of an entire force which must be examined if the effectiveness of the force is to be determined. The total force thus includes not only maneuver elements, but also ground and air support elements.

2. A force may exhibit mobility without respect to its mission. A successful attack, defense, delay, or retrograde is highly dependent on varying levels of a force's mobility.

3. The operational mobility of a force is relative, *i.e.*, an advantage may be gained by denying the opposing force some portion of his mobility.

4. The speed and size of a force are critical indicators of mobility potential and may prove advantageous or detrimental depending upon their tactical or strategic use.

5. In an evaluation of operational effectiveness, operational mobility must be considered in concert with other force attributes such as type and status of equipment, firepower, leadership, morale, and *esprit de corps*.

A suitable definition of operational mobility based on the above historical analysis, then, would necessarily include measures of the relative speeds and sizes of opposing forces with appropriate considerations given to tactical and strategical employment in any given combat environment for a variety of combat missions. Furthermore, these characteristics must be conformable to quantification so that operational mobility may be examined in concert with other force attributes, such as firepower. The authors of the RAC study propose such a measurement in terms of force momentum, where momentum is physically expressed as the

product of mass and velocity (2). A more detailed examination of this proposal will be presented in a subsequent section.

This macro-view of operational mobility is predicated on its recognition as a force attribute considered in the same plane as other force attributes. Focus will now be directed toward a concept that also considers operational mobility as a force attribute, but represents it as a combination of system-oriented attributes.

Micro-view of Operational Mobility

Major research efforts to date imply a recognition of the fact that performance characteristics of land combat vehicles influence the operational performance of the force in which they are employed (42). Thus, if an APC has a limited ability to traverse battlefield obstacles (both natural and man-made), then throughout the course of a battle a force's maneuverability may be hindered. Similarly, logistical support may be hampered by restricted or limited carrying capacity of prime-mover vehicles. This recognition has led to the development of a highly sophisticated mobility model in an effort to predict the performance of a vehicle over specific types of terrain, given specific performance characteristics (24). Predicted performance data is then used as input data for stochastic or deterministic simulations of land combat. A method for determining how these performance characteristics influence the operational mobility of a force has been proposed in (42). Essentially, this method compares the force effectiveness MOE obtained from a land combat simulation utilizing the performance input of one vehicular system to the MOE obtained utilizing the performance input of another vehicular system. Differences in force effectiveness (all other input parameters being held

constant) are then attributed to the mobility differentials of the two systems.

Although this procedure provides an intuitively appealing approach to the problem, it circumvents the issue of determining a precise measure of operational mobility and its exact contribution to force effectiveness. However, if the effect of a finite number of individual performance characteristics on the mobility of a force may be determined, then individual and overall contributions to force effectiveness may also be determined.

Consider the ability of a land combat vehicle to traverse battlefield obstacles. A measure of this performance characteristic in a simulated battle may be the number of times that vehicles of a given type fail to traverse any obstacle. Similar measures of performance in terms of speed may be determined, *e.g.*, the number of times vehicles of a particular type delay ordered troop movements or fail to keep pace with other vehicles in tactical formations. For a given set of performance parameters, then, there are a corresponding number of mobility performance measures that completely describe the performance of a particular type vehicle in terms of its effect on the operational mobility of the force. The operational mobility of a force may thus be described and measured in terms of a vehicular system's operational performance measures, where the performance measures are considered as a completely descriptive subset of operational mobility. Theoretically, there would be a subset of performance measures for each type vehicle in the force. A certain amount of subjective judgment remains due to the fact that a decision must be made with respect to which performance measures will be

considered to completely describe the operational mobility of the force for each type vehicle. This recognized limitation of the micro-concept and its impact on the proposed methodology will be discussed in Chapter 6.

Proposed Definitions

Based on the preceding discussion, the following two definitions of operational mobility are proposed:

1. The operational mobility of a force is the momentum of that force measured with respect to the momentum of an opposing force, *i. e.*, relative momentum.
2. The operational mobility of a force is the collection of subsets of operational performance measures obtained for specified vehicular performances characteristics.*

Measures of individual vehicle performance characteristics have been proposed and examined in (42); hence, no detailed justification for their use will be proposed in this study. However, since there is little information available regarding the concept of relative force momentum as a measure of operational mobility, the exact measurement of this attribute will now be examined and a military-based justification provided.

Measurement of Relative Momentum

From any text in elementary physics, for example (15), the following mathematical expression for the momentum, p , of a solid body may be obtained:

*This definition is also conformable to relative measure as the ratio of or difference between corresponding performance measures obtained for opposing forces.

$$p = mv ,$$

where m is the mass of the body and v is its velocity. Furthermore, mass is expressed as the quotient of the weight of the body and gravitational force, or

$$m = \frac{w}{g} .$$

The application of this concept to the analysis of the effectiveness of military forces through the use of simulated combat models necessitates the examination of several physical notions.

1. Mass of military forces:

The forces must be considered as "solid" bodies in order to apply principles of momentum. This assumption is realistic, however, since even solid bodies in physics contain small spaces in their molecular structure. Taken in a military context, then, these spaces may be equated to distances between elements, or even between vehicles and personnel. In light of this, the mass of a military force is merely its total weight (weapons, vehicles, personnel, equipment, etc.) divided by g , the force of gravity. As attrition occurs throughout the course of a simulated battle, the mass of the force will be reduced.

2. Velocity of a military force:

The velocity, v , of a force may be expressed as:

$$v = \text{distance traveled} \div \text{time to travel that distance}$$

An apparent drawback may be recognized when considering forces in a de-

fensive posture (*i.e.*, $v = 0$). This would not affect a measure of momentum, however, if the force's velocity were measured with respect to the centroid (center of mass) of the force. In this way, even "stationary" forces would exhibit velocity as attrition occurs, reserves are committed, and so forth.

3. Tactical impact:

The *concept* of momentum as a definition of operational mobility is not influenced by tactical decisions, mission of a force, terrain, or any other military considerations. Different situations will merely result in obtaining different numerical values for the momentum measure. For example, if a force loses half of its assigned APC's due to faulty route selection by the unit commander, these losses will be reflected in MOE describing personnel and vehicle losses; similarly, these losses may be reflected by reduced momentum of the force since the total weight of the force will have decreased.

4. Measurement:

The momentum of a force may be computed at any specified time throughout a simulated battle by the relationship presented earlier and averaged over the entire battle to obtain a measurement of average momentum. Since the momentum of each of the opposing forces may be computed, relative momentum may be expressed as a ratio in a manner equivalent to that traditionally used to compute loss exchange ratios (LER) for personnel and equipment losses (48).

Relative Momentum as a MOE to Describe Operational Mobility

Justification for defining operational mobility in terms of relative momentum will be based primarily on the principles of war. The

principle of mass dictates that "superior combat power must be concentrated at the critical time and place for a decisive purpose" (45).

Degrees of adherence to this principle will be directly reflected by the representation of mass in the equation for momentum. As mentioned previously, mission, terrain, and tactical decisions with respect to time and place of troop concentrations will also be reflected merely as different numerical values for momentum. Thus, a direct analogy may be drawn between the mass of a force and the principle of mass.

Speed, or velocity, of a force may be analyzed in terms of several principles of war (quotations below are all from reference 45):

1. Objective: "The objective of each operation must contribute to the ultimate objective", *i. e.*, the defeat of enemy forces. "Each intermediate objective must be such that its attainment will most directly, quickly, and economically contribute to the purpose of the operation".

2. Economy of force: "Minimum essential means must be employed at points other than that of the main effort. This principle is the reciprocal of the principle of mass".

3. Surprise: "Surprise results from striking an enemy at a time and place and in a manner for which he is unprepared.... Factors contributing to surprise include speed, cover and deception...."

Indirect analogies may also be drawn between the speed of a force and other principles of war, *e. g.*, the principle of offensive, which dictates seizing the initiative through exploitation of enemy weaknesses. The direct analogies presented above, however, sufficiently portray the potential importance of speed and mass on the battlefield. A measure of effectiveness that encompasses all but a few principles of war, then,

will necessarily contain a large amount of information about the effectiveness of a force engaged in combat.

Structure of the Problem

Considering MOE in terms of the two proposed definitions, the problem of determining the contribution of these and other MOE to the operational effectiveness of a force begins to assume an identifiable structure. With respect to Definition 1, relative momentum may be considered in the same plane as other MOE (*e. g.*, LER) to describe operational effectiveness. Hence,

$$\underline{OE'} = [(\underline{LER})_1, (\underline{LER})_2, \dots, \underline{M}] ,$$

where (') denotes the transpose of the column representation of OE and each LER and the measure of relative momentum (M) is represented by a set of observations obtained from replications of simulated combat. Conforming to standard notation to be used throughout the remainder of this paper, the above situation may be described by the following matrix of observations:

$$\underline{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kn} \end{bmatrix} , \quad (2.1)$$

where each of the n columns represents a set of k observations on each of the MOE being examined.

Examination of operational effectiveness using Definition 2 con-

forms to a similar structure:

$$OE' = [(LER)_{\tilde{1}}, (LER)_{\tilde{2}}, \dots, M_{\tilde{1}}, M_{\tilde{2}}, \dots, M_{\tilde{r}}] ,$$

where each measure of operational mobility (M) is now a set of operational performance measures tabulated for each of the r vehicles being considered. Representation of this description yields a data matrix similar to (2.1), the only difference possibly being the dimensions of the matrices.

Since each of the $X_{\tilde{i}}$ (MOE) in the data matrices is a random variable and is quite possibly correlated with other $X_{\tilde{i}}$, exploitation of possible correlations must be undertaken in order to obtain precise information concerning those MOE being examined. This suggests application of multivariate statistical methods, which will be reviewed in the next chapter concurrently with the development of a general methodology for accomplishing the stated objective.

CHAPTER III

REVIEW OF APPLICABLE MULTIVARIATE STATISTICAL METHODS

Introduction

The purpose of this chapter is to review those multivariate statistical methods that are applicable to the objective of this research. Rather than discussing these methods in an isolated statistical vacuum, however, their relationship to the methodological questions at hand will be established so that a general methodology will become apparent as the discussion progresses. In this way, formulation of the precise methodology to be presented in Chapter 4 will be merely a synthesis of methods presented in this chapter.

No attempt will be made to trace the detailed and rigorous theoretical development of multivariate statistics; however, some theoretical discussion will be necessary in order to establish the relationship between the well-documented methods examined and the general methodology.

Major areas to be discussed include an examination of the assumption of multivariate normality, hypothesis testing, simultaneous confidence intervals, and principal component analysis. Before proceeding, however, it will prove beneficial to survey several basic multivariate statistical concepts and to preview the notation to be used throughout the remainder of this paper. Most of the notation conforms to that used in basic textbooks on multivariate statistics, *e. g.*, (5) and (30).

Basic Multivariate Statistical Concepts and Notation

In examining a p -dimensional vector of continuous random variables as introduced in Chapter II and represented by:

$$\underline{X}' = [X_1, X_2, \dots, X_p],$$

several basic relationships become important. The covariance between any two elements of the population of \underline{X} , say, X_i and X_j , is defined as:

$$\text{Cov } (X_i, X_j) = E \{ [X_i - E(X_i)] [X_j - E(X_j)] \}, \quad (3.1)$$

where E is the usual expected value operator. If $i = j$, then (3.1) reduces to:

$$\text{Cov } (X_i, X_i) = \text{Var } (X_i),$$

the variance of the i th element of \underline{X} (18). Furthermore,

$$\text{Cov } (X_i, X_j) = \text{Cov } (X_j, X_i).$$

The $p \times p$ symmetric matrix whose elements are the population variances and covariances of \underline{X} is termed the population covariance matrix and is represented by:

$$\Sigma_X = \begin{bmatrix} \text{Var } (X_1) & \text{Cov } (X_1, X_2) & \dots & \text{Cov } (X_1, X_p) \\ \text{Cov } (X_2, X_1) & \text{Var } (X_2) & \dots & \text{Cov } (X_2, X_p) \\ \vdots & \vdots & \dots & \vdots \\ \text{Cov } (X_p, X_1) & \text{Cov } (X_p, X_2) & \dots & \text{Var } (X_p) \end{bmatrix} \quad (3.2)$$

If $\underline{x}' \Sigma_X \underline{x} > 0$ for all non-null \underline{x} , then Σ_X is said to be positive definite (17). In a similar manner, the correlation between any two ele-

ments of \tilde{X} is defined by the Pearson correlation coefficient in (17) as

$$\rho_{ij} = \frac{\text{Cov}(\tilde{X}_i, \tilde{X}_j)}{\sqrt{\text{Var}(\tilde{X}_i)} \cdot \sqrt{\text{Var}(\tilde{X}_j)}} \quad (3.3)$$

and the matrix of correlations between elements of \tilde{X} is:

$$\tilde{R} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}, \quad (3.4)$$

where $\rho_{ij} = \rho_{ji}$ and $\rho_{ii} = 1$ for all $i \leq p$. If \tilde{X}_i and \tilde{X}_j are independent, then $\rho_{ij} = 0$. However, the converse is not always true; thus, if $-1 \leq \rho_{ij} \leq 0$ or $0 \leq \rho_{ij} \leq 1$ for $i \neq j$, then \tilde{X}_i and \tilde{X}_j are said to be dependent. The vector of population means of the elements of \tilde{X} will be denoted by:

$$\mu_{\tilde{X}}' = [\mu_{X1}, \mu_{X2}, \dots, \mu_{Xp}], \text{ where}$$

μ_{Xi} is the mean of \tilde{X}_i .

From a data matrix representing N observations on the elements of \tilde{X} , such as that represented by equation (2.1), the p by p symmetric matrix of sums of squares and cross products may be calculated:

$$A = \sum_{i=1}^N (\tilde{X}_i - \bar{\tilde{x}})(\tilde{X}_i - \bar{\tilde{x}})', \quad (3.5)$$

where $\bar{\tilde{x}}$ is the vector whose elements are the maximum likelihood estimates of the elements of $\mu_{\tilde{X}}$. The maximum likelihood estimator of the popula-

tion covariance matrix, or the sample covariance matrix, may thus be calculated by:

$$\underline{S}_X = \frac{1}{N - 1} A. \quad (3.6)$$

Although the preceding discussion by no means represents all concepts and notation to be used in the sequel, sufficient groundwork has been established so that continuing with the examination of multivariate statistical methods is now possible.

The Assumption of Multivariate Normality

As in univariate statistics, the distributional theory for multi-dimensional, continuous random variables is well-documented for the multivariate analog of the univariate normal distribution. Although exact distributions for other multivariate statistics have been derived (8), only statistics based on multivariate normality of a vector random variable will be discussed based on the previously mentioned documentation and universality of use.

In an examination of the vector random variable, \underline{X} , then, it would be extremely beneficial if the joint density function of \underline{X} , $\phi(\underline{X})$, were of the form:

$$\phi(\underline{X}) = \frac{1}{(2\pi)^{p/2} |\underline{\Sigma}_X|^{1/2}} \exp \left[-1/2 (\underline{X} - \underline{\mu})' \underline{\Sigma}_X^{-1} (\underline{X} - \underline{\mu}) \right], \quad (3.7)$$

where $|\underline{\Sigma}_X|$ and $\underline{\Sigma}_X^{-1}$ are the determinant and inverse of the sigma-matrix, $\underline{\Sigma}_X$, respectively. If the joint density of \underline{X} is of the form of (3.7), then \underline{X} is said to follow a multivariate normal distribution, or

$\tilde{X} \sim N(\mu_{\tilde{X}}, \Sigma_{\tilde{X}})$. Furthermore, if $\Sigma_{\tilde{X}}$ is of full rank p , then $|\Sigma_{\tilde{X}}| \neq 0$, $\Sigma_{\tilde{X}}^{-1}$ always exists, and $\Sigma_{\tilde{X}}$ is positive definite (1).

Therefore, in a statistical analysis of \tilde{X} , several alternatives exist:

1. Assume $\tilde{X} \sim N(\mu_{\tilde{X}}, \Sigma_{\tilde{X}})$ and proceed with the analysis;

2. Test $H_0 : \tilde{X} \sim N(\mu_{\tilde{X}}, \Sigma_{\tilde{X}})$ against

$H_1 : \tilde{X}$ is not $N(\mu_{\tilde{X}}, \Sigma_{\tilde{X}})$ and

proceed with the analysis if H_0 is not rejected; or

3. Transform \tilde{X} by $\tilde{Y} = f(\tilde{X})$ so that $\tilde{Y} \sim N(\mu_{\tilde{Y}}, \Sigma_{\tilde{Y}})$ and such that inferences with respect to \tilde{X} may be related to inferences drawn on \tilde{Y} .

4. Disregard the distribution of \tilde{X} and proceed with a nonparametric analysis.

For large samples drawn from the population of \tilde{X} , alternative 1 is certainly tenable based on a multivariate generalization of the Central Limit Theorem (5). In most analyses of military operational effectiveness, however, obtaining large sample sizes (*e. g.*, >50) is usually not feasible due to costs associated with replications of stochastic combat simulation models. Alternative 4 is currently being examined by an Army agency and will not be discussed in this paper. This leaves alternatives 2 and 3, or a combination thereof, as possible approaches in commencing an analysis of \tilde{X} .

Tests for Multivariate Normality

Until recently this area of multivariate statistics has received little attention. However, generalizations of univariate test procedures have been formulated. One such procedure, based on order statistics, is due to Malkovich and Afifi (27) and is a generalization of a univariate

procedure developed by Shapiro and Wilk (38). Given the data matrix $\tilde{X} = [x_{ij}]$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, n$ and computing the matrix of sums of squares and cross products, A , as in equation (3.5), the test proceeds as follows:

1. Denote \tilde{X}_i as the $p \times 1$ observation vector for which

$$(\tilde{X}_i - \bar{x})' A^{-1} (\tilde{X}_i - \bar{x}) = \max_{1 \leq j \leq n} (\tilde{X}_j - \bar{x})' A^{-1} (\tilde{X}_j - \bar{x}).$$

2. Order the statistics

$$U_j = (\tilde{X}_i - \bar{x})' A^{-1} (\tilde{X}_j - \bar{x}), \quad j = 1, 2, \dots, n \quad (3.8)$$

and denote them by U_1, U_2, \dots, U_n .

3. Compute the test statistic, W :

$$W = \frac{\left[\sum_{i=1}^n a_i U_i \right]^2}{(\tilde{X}_i - \bar{x})' A^{-1} (\tilde{X}_i - \bar{x})} \quad (3.9)$$

where the $\{a_i\}$ are symmetric, normalized best linear coefficients tabulated in (36) and presented in Table 20, Appendix A. Exact values of $\{a_i\}$ for sample sizes up to 20 are known, while values for sample sizes up to 50 have been approximated in (38).

Malkovich and Afifi compared this test with others based on multivariate generalizations of definitions of univariate skewness and kurtosis and generalizations of the univariate Kolmogorov-Smirnoff and Cramer-Von Mises test statistics. The comparison examined Monte Carlo

approximations of the powers of the respective tests against selected alternative distributions and showed that in most instances with small sample sizes (< 20), the W statistics performed as well as, if not better than, the other tests. These results, and the fact that the exact distribution of W is known for sample sizes up to 50 (Table 21, Appendix A), combine with the relative ease of the procedural computation of W to make the test appealing for practical application. Shapiro and Wilk, in the presentation of their univariate procedure, indicate that the W-statistic is somewhat sensitive to outliers, which may be verified by a cursory examination of equation (3.9). This characteristic, however, may serve to the analyst's advantage by detection of even slight departures from normality, unless outliers are due to random error. In light of this, a relatively low confidence level associated with a given W, say $\alpha = .7$ or $\alpha = .6$, may be suitable if the presence of outliers may be verified.

Other tests for multivariate normality have been proposed, *e. g.*, (26, 28, 29); however, computational difficulties and lack of knowledge of exact distributions of related test statistics for \tilde{X} of dimension $p > 2$ prohibit their use for the problem situation of this paper, in which the number of MOE being examined will most certainly be more than two.

Data Transformation

In an examination of alternative 3, several items become immediately important. If the transformation $\tilde{Y} = f(\tilde{X})$ is of the form $\tilde{Y} = C\tilde{X} + \tilde{b}$, *i. e.*, a linear transformation, then inferences concerning

\tilde{X} are equivalent to those drawn with respect to \tilde{Y} (33). However, no linear transformation is known to exist that *simultaneously* induces joint normality on all p dimensions of \tilde{X} . Each element of \tilde{X} may be transformed separately to insure marginal normality of each X_{i1} , but this is not sufficient to make any inferences about the joint distribution of \tilde{X} . An examination of a non-linear transformation of \tilde{X} may then prove useful.

Development of transformations to insure multivariate normality has proceeded much in the same direction as the development of tests for multivariate normality through generalization of univariate procedures. Box and Cox (12) presented an extensive examination of several univariate procedures, one of which was based on the class of power transformations

$$\tilde{Y} = \begin{cases} \frac{X_i^{\lambda-1}}{\lambda} & \lambda \neq 0 \\ \ln(X_i) & \lambda = 0 \end{cases},$$

where \tilde{Y} is now the transformed vector of original observations, $\{X_i\}$, following the univariate normal distribution. If \tilde{Y} is assumed to be normally distributed for some unknown parameter, λ , then maximization of the log likelihood function, $L_{\max}(\lambda)$, will produce that value of λ that makes \tilde{Y} normally distributed.

This procedure has been generalized in (3) by Andrews, Gnanadesikan, and Warner as a method for insuring joint normality. Given the original data matrix $\tilde{X} = [x_{ij}]$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, n$ and a starting vector of transformation parameters $\tilde{\lambda}' = [\lambda_1, \lambda_2, \dots, \lambda_p]$, proceed as follows:

1. Calculate

$$\tilde{Y}^{(\lambda)} = y_{ij}^{(\lambda)} = \begin{cases} \frac{x_{ij}^{\lambda_j} - 1}{\lambda_j}, & \lambda_j \neq 0 \\ \ln(x_{ij}), & \lambda_j = 0 \end{cases} \quad (3.10)$$

2. Using $\tilde{Y}^{(\lambda)}$, the transformed data matrix, calculate the maximum likelihood estimates of the mean vector and the covariance matrix by:

$$\begin{aligned} \bar{\tilde{y}} &= \frac{1}{n} \tilde{Y}^{(\lambda)} \cdot \underline{1} \\ \underline{S} &= \frac{1}{n} (\tilde{Y}^{(\lambda)} - \underline{1} \cdot \bar{\tilde{y}}')' (\tilde{Y}^{(\lambda)} - \underline{1} \cdot \bar{\tilde{y}}') \end{aligned}$$

where $\underline{1}$ is a p -dimensional vector of one's.

3. Numerically maximize

$$L_{\max}(\hat{\lambda}) = -\frac{n}{2} \ln |\underline{S}| + \left(\sum_{i=1}^p (\hat{\lambda}_i - 1) \sum_{j=1}^n \ln(x_{ij}) \right) \quad (3.11)$$

to obtain $\hat{\lambda}$, the vector of maximum likelihood estimates of the transformation parameters.

Theoretically, the only set of parameters, λ , that is consistent with the hypothesis that $\underline{X} \sim N(\underline{\mu}_X, \underline{\Sigma}_X)$ is $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_p) = \underline{1}$ (3). Based on this and an extension of the results in (12), a $100(1-\alpha)$ confidence level may be established for $\hat{\lambda}$, where the test statistic

$$2[L_{\max}(\hat{\lambda}) - L_{\max}(\underline{1})] \sim \chi_p^2 \quad (3.12)$$

when performing a likelihood ratio test of $H_0: \lambda = \underline{1}$ (3).

Also, since $-\infty < \hat{\lambda} < \infty$, a set of transformation parameters that has all elements relatively close to 1, say $.9 < \hat{\lambda}_i < 1.1$, $i = 1, \dots, p$, indicates that the original set of data did not have to be subjected to a large disturbance in order to enhance its normality (3). This transformation procedure, then, actually provides several indicators in an assessment of multivariate normality.

In comparing this method with several others, the authors of (3) concluded that the procedure outlined above performed as well as the others against similar alternative hypotheses. This procedure is more practical, however, because of the exact nature of its derivation and relative ease of computation. Other procedures developed were based on approximate directional and angular perturbances of the data and were generally not as precise as the maximum likelihood procedure.

The only physical restriction on the above procedure is that $x_{ij} > 0$ and $y_{ij}^{(\lambda)} > 0$, since the natural logarithm of a negative number is not defined. If necessary, the means of \tilde{X} may be suitably shifted at the outset of an analysis to insure satisfaction of this criteria. From basic statistics, shifts of this nature will not affect the underlying distributions of \tilde{X} nor of \tilde{Y} . Attention will be directed in a later section to the issue of drawing inferences on the original variable, \tilde{X} , after application of a non-linear transformation.

Tests of Hypotheses

The purpose of this section is to introduce certain statistical tests that may prove useful in a preliminary examination of MOE being considered. Precise contributions of each MOE to operational effective-

ness can not be determined simply by testing the structure of a system of data; however, approximate inferences may be drawn which may preclude the necessity for further testing and thus reduce associated costs. Numerous statistical tests are available which might prove advantageous (22); however, only well-documented tests will be examined.

Testing Equality of Covariance Matrices

In most operational effectiveness analyses, of primary concern is the examination of several competing systems so that the effectiveness of a force with one system may be compared against the effectiveness of a force with a base system or another alternative system. With respect to output from a stochastic simulation, then, the concern is with examination of the data structure of each system. As in univariate hypothesis testing, first consideration should be given to testing equality of the covariance matrices for competing systems, which is simply the multivariate analog of testing equality of variances in the univariate case (homoscedasticity).

If only two systems are examined at one time, say the base system and the first alternative, then the hypothesis to be tested is:

$$H_0: \Sigma_{X1} = \Sigma_{X2}$$

$$\text{against } H_1: \Sigma_{X1} \neq \Sigma_{X2}$$

where Σ_{X1} has dimension p_1 and Σ_{X2} has dimension p_2 (both of full rank). Morrison (30) provides a generalization of a univariate test developed by Bartlett based on the generalized likelihood-ratio criterion. If the observations for each system are drawn from a multivariate normal popu-

lation, then the estimates of Σ_{X1} and Σ_{X2} , S_1 and S_2 , may be computed in the usual manner using equation (3.6), where N is replaced by N_1 and N_2 respectively. Letting $n_i = N_i - 1$, the pooled estimate of Σ , S_p , when H_0 is true becomes:

$$S_p = \sum_{i=1}^2 \left(\frac{n_i S_i}{\sum_{i=1}^2 n_i} \right), \quad (3.13)$$

and the statistic is:

$$M = \left(\sum_{i=1}^2 n_i \right) \ln |S_p| - \sum_{i=1}^2 n_i \ln |S_i|. \quad (3.14)$$

If k is the number of systems being tested, Box (11) has shown that introduction of the scale factor

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \sum_{i=1}^k \frac{1}{n_i} - \frac{1}{\sum_{i=1}^k n_i} \quad (3.15)$$

insures that

$$MC^{-1} \sim \chi^2_{1/2(k-1)p(p+1)} \quad (3.16)$$

for k and $p \leq 4$ or 5 and each $n_i > 20$. For larger dimensions and sample sizes, Box proposes

$$Mb^{-1} \sim F_{f_1, f_2} \quad (3.17)$$

for equal sample sizes, $N_1 = N_2 = N$, and where

$$f_1 = p_1 p_2, \quad f_2 = \frac{12(N-1)^2 (p_1 p_2 + 2)}{p_1^2 + p_2^2 - 5}$$

and

$$b^{-1} = 1 - \frac{p_1 + p_2 + 1}{2(N-1)} - \frac{f_1}{f_2} \cdot \frac{p_1 p_2}{p_1 p_2} \quad (3.18)$$

If $MC^{-1} \leq \chi^2_{1-\alpha; 1/2(k-1)p(p+1)}$ or $Mb^{-1} \leq F_{1-\alpha; f_1, f_2}$, the null hypothesis is not rejected with confidence 100 (1- α)%.

Tests of Hypotheses on Mean Vectors

Several tests on mean vectors of data obtained for two competing systems will prove useful. If statistical inferences indicate no differences between the means of two systems, then further statistical analysis may not be warranted. Similarly, established statistical differences may justify further analysis and may in fact indicate which elements are different and may therefore contribute different amounts to operational effectiveness.

If $H_0: \Sigma_{X1} = \Sigma_{X2}$ is not rejected then a test of whether or not the means of two systems are different may be stated as:

$$H_0: \mu_{X1} = \mu_{X2}$$

$$\text{against } H_1: \mu_{X1} \neq \mu_{X2}.$$

If $X_1 \sim N(\mu_{X1}, \Sigma_{X1})$ and $X_2 \sim N(\mu_{X2}, \Sigma_{X2})$ and the true population covariance matrices are unknown, Morrison (30) provides the following statistic:

$$F = \frac{N_1 + N_2 - p - 1}{(N_1 + N_2 - 2)p} T^2 \sim F_{p, N_1 + N_2 - p - 1} \quad (3.19)$$

where

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{S}_p^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad (3.20)$$

is the Hotelling T^2 statistic (21), the multivariate analog of the squared univariate t - statistic. In equation (3.20), \mathbf{S}_p is the usual pooled estimate of Σ , while $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ are maximum likelihood estimates of μ_{X1} and μ_{X2} . If

$$T^2 \leq \frac{(N_1 + N_2 - 2)p}{(N_1 + N_2 - p - 1)} F_{\alpha; p, N_1 + N_2 - p - 1} = T_0^2, \quad (3.21)$$

then H_0 is not rejected with confidence 100 (1- α)%.

Another useful test may be

$$\begin{aligned} H_0: \mu_{Xi} &= \mu_0 \\ \text{against } H_1: \mu_{Xi} &\neq \mu_0 \end{aligned}$$

where μ_0 is a given vector of pre-established standard means, Σ_{X1} are unknown, and $\mathbf{X}_1 \sim N(\mu_{X1}, \Sigma_{X1})$. When H_0 is true, Morrison shows that

$$F = \frac{N - P}{p(N - 1)} T^2 \sim F_{p, N - p}, \quad (3.22)$$

where T^2 is now defined as

$$T^2 = N(\bar{x}_i - \mu_0)' S_i^{-1} (\bar{x}_i - \mu_0) . \quad (3.23)$$

If

$$T^2 \leq \frac{p(N-1)}{N-p} F_{\alpha; p, N-p} = T_0^2 , \quad (3.24)$$

then H_0 is not rejected with confidence $100(1-\alpha)\%$.

When $H_0: \Sigma_{X1} = \Sigma_{X2}$ is rejected, a different procedure must be used to test $H_0: \mu_{X1} = \mu_{X2}$. The situation now encountered is the multivariate analog of the Behrens-Fisher problem. A solution to this problem has been derived by Bennett (7) based on an extension of Sheffé's solution to the univariate problem (40) and presented in Anderson (5). Only the case of equal sample sizes will be considered, *i. e.*, $N_1 = N_2$.

By forming the matrix of differences between observations on X_1 and X_2

$$Y = \{y_{ij}\} = \{X_{1j} - X_{2j}\} ,$$

the following quantities may be calculated:

$$\bar{y} = \frac{1}{N} Y' \cdot \underline{1} \quad (3.25)$$

and

$$S = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})(y_i - \bar{y})' . \quad (3.26)$$

Accordingly,

$$T^2 = N\bar{y}' S^{-1} \bar{y} \quad (3.27)$$

has the T^2 distribution with $N - 1$ degrees of freedom. If

$$T^2 \leq \frac{P(N - 1)}{N - p} F_{\alpha; p, N - p} = T_0^2 \quad (3.28)$$

then H_0 is not rejected with confidence $100(1-\alpha)\%$. This problem does not arise when testing $H_0: \mu_{xi} = \mu_0$ since no estimate of a pooled covariance matrix is needed to establish an appropriate critical region.

As mentioned previously, if no differences in means are detected in tests outlined above, further testing may be unnecessary. However, if a null hypothesis on the mean vectors of two systems is rejected, then it is desirable to know which elements of the mean vectors, *i. e.*, which MOE, are contributing to differences in operational effectiveness. This may in fact be determined and is the topic of the next section.

Simultaneous Confidence Intervals

Once again, procedures reviewed here are direct extensions of univariate procedures as presented in (5) and (30), unless otherwise stated. By establishing confidence intervals about the elements of mean vectors and/or differences between mean vectors, it will be possible to determine exactly which elements caused rejection of a particular null hypothesis.

Consider first $H_0: \mu_{x1} = \mu_{x2}$ where $H_0: \Sigma_{x1} = \Sigma_{x2}$ is not rejected. The $100(1-\alpha)\%$ simultaneous confidence intervals on $a'(\mu_{x1} - \mu_{x2})$ are

$$a'(\bar{x}_1 - \bar{x}_2) \pm \left[a' S_p a \frac{(N_1 + N_2)(N_1 + N_2 - 2)p}{(N_1 N_2)(N_1 + N_2 - p - 1)} F_{\alpha; p, N_1 + N_2 - p - 1} \right]^{\frac{1}{2}}, \quad (3.29)$$

where a' is a vector of zero's except for 1 in the i th position corresponding to the i th interval.

If $H_0: \Sigma_{X1} = \Sigma_{X2}$ is rejected, then 100(1- α)% simultaneous confidence intervals on $a'(\mu_{X1} - \mu_{X2})$ are constructed in a similar manner, based on results of the Behrens-Fisher problem:

$$a'(\bar{x}_1 - \bar{x}_2) \pm \left[a' S_y a \frac{(N - 1)}{N(N - p)} F_{\alpha; p, N - p} \right]^{\frac{1}{2}}. \quad (3.30)$$

For establishing intervals about $a'(\mu_{X1} - \mu_0)$, equation (3.30) may also be used by replacing $(\bar{x}_1 - \bar{x}_2)$ with $(\bar{x}_1 - \mu_0)$ and S with S_1 from equation (3.6).

Interpretation of intervals established by any of the above procedures is critical. Consider an example in a test of $H_0: \mu_{X1} = \mu_{X2}$ where X_1 and X_2 are both of dimension $p = 2$ and H_0 has been rejected. Intervals established by equation (3.30) are found to be:

$$L_1 \leq \mu_{X11} - \mu_{X12} \leq U_1$$

and $L_2 \leq \mu_{X21} - \mu_{X22} \leq U_2$

If the closed interval $[L_1, U_1]$ contains 0, then $\mu_{X12} - \mu_{X22}$ is considered negligible at the specified confidence level and did not contribute significantly to the rejection of H_0 . If $[L_1, U_1]$ does not contain 0,

then the converse is true. Hence, statistical differences may be established between specific MOE used to describe operational effectiveness, provided that differences between MOE of competing systems have been established. If a non-linear transformation is applied to \underline{X}_i to enhance multivariate normality, however, additional considerations must be examined with respect to inferences drawn on the transformed data.

Statistical Inference After Non-Linear Transformation

In general, the situation now encountered may be separated into two cases: 1) when testing $H_0: \mu_{X1} - \mu_{X2} = 0$; and 2) when testing $H_0: \mu_{X1} - \mu_0 = 0$.

Considering Case 1, confidence interval estimates of the means of the transformed data are of the form

$$L_j \leq \mu_{Y1j} - \mu_{Y2j} \leq U_j, \quad (3.31)$$

where μ_{Y1j} represents the mean of the j th transformed variable in data set i , transformed by $Y_i = f_i(X_i)$ from equation (3.10). Since Y is non-linear, the intervals $[L_j, U_j]$ must be reconstructed to be of the form

$$g(L_j) \leq \mu_{X1j} - \mu_{X2j} \leq g(U_j),$$

where g is some function of $[L_j, U_j]$ such that inferences drawn from $g[L_j, U_j]$ are equivalent to, but not necessarily the same as, those drawn from $[L_j, U_j]$. Several alternatives to the solution of this situation become readily apparent.

First, the exact distributions of the transformed parameters may be derived so that statistics based on these distributions yield the re-

quired interval estimates. This would actually preclude the necessity for reconstruction of intervals on the transformed parameters, but would dictate derivation of exact tests for the transformation used.

A second method would be to develop an inverse function of f_i such that inferences drawn from $f_i[L_j, U_j]$ are equivalent to those drawn from $[L_i, U_i]$. Computational burdens reduce the practicality of adopting this approach.

A third alternative might be an attempt at the identification of a realistic, intuitive interpretation of the transformed interval estimates in terms of the original data. For example, if $y = f(x) = \ln x$, then y may be interpreted as a measurable variable representing some growth pattern of x . Although this procedure has proved useful in some instances (23), its application to military problems involves substantial conjecture and subjective evaluation.

The first two procedures involve statistical techniques and methods of functional analysis beyond the scope of this research, while the third procedure is not readily applicable to the problem at hand. In light of this, an approximate method will be developed to obtain required interval estimates of the original parameters, $\mu_{X1j} - \mu_{X2j}$.

Based on the independence of the transformed data sets, Y_i , $\mu_{Y1j} - \mu_{Y2j} = E(Y_{1j}) - E(Y_{2j})$. Recalling that the transformation is

$$Y_{ij} = \frac{X_{ij}^{\hat{\lambda}_j} - 1}{\hat{\lambda}_j}$$

for $\hat{\lambda} \neq 0$, $j \leq p$, approximations for $E(Y_{1j})$ and $E(Y_{2j})$ may be obtained

from (18),

$$E(Y_{ij}) = f_i(\mu_{Xij}) + 1/2 f_i''(\mu_{Xij}) \sigma_{ij}^2 + R_j, \quad (3.32)$$

based on the Taylor series expansion of f_i about $X_{ij} = \mu_{Xij}$.

Using only the first order approximation, equation (3.31) may be written as

$$L_j \leq E(Y_{1j}) - E(Y_{2j}) \leq U_j, \quad \text{or} \quad (3.33)$$

$$L_j \leq f_1(\mu_{X1j}) - f_2(\mu_{X2j}) \leq U_j.$$

Substituting further gives

$$L_j \leq \frac{\mu_{X1j}^{\hat{\lambda}_{1j}-1}}{\hat{\lambda}_{1j}} - \frac{\mu_{X2j}^{\hat{\lambda}_{2j}-1}}{\hat{\lambda}_{2j}} \leq U_j.$$

Combining denominators gives

$$L_j \leq \frac{\hat{\lambda}_{2j}(\mu_{X1j}^{\hat{\lambda}_{1j}-1}) - \hat{\lambda}_{1j}(\mu_{X2j}^{\hat{\lambda}_{2j}-1})}{\hat{\lambda}_{1j}\hat{\lambda}_{2j}} \leq U_j,$$

or

$$\hat{\lambda}_{1j}\hat{\lambda}_{2j}L_j \leq \hat{\lambda}_{2j}\mu_{X1j}^{\hat{\lambda}_{1j}} - \hat{\lambda}_{2j} - \hat{\lambda}_{1j}\mu_{X2j}^{\hat{\lambda}_{2j}} + \hat{\lambda}_{1j} \leq U_j.$$

So,

$$\hat{\lambda}_{2j} - \hat{\lambda}_{1j} + \hat{\lambda}_{1j}\hat{\lambda}_{2j}L_j \leq \hat{\lambda}_{2j}\mu_{X1j}^{\hat{\lambda}_{1j}} - \hat{\lambda}_{1j}\mu_{X2j}^{\hat{\lambda}_{2j}} \leq \hat{\lambda}_{2j} - \hat{\lambda}_{1j} + \hat{\lambda}_{1j}\hat{\lambda}_{2j}U_j.$$

Denoting the left and right hand sides of the above inequality by L'_j and U'_j , respectively, gives

$$L'_j \leq \hat{\lambda}_{2j}^{\mu_{X1j}} \hat{\lambda}_{1j} - \hat{\lambda}_{1j}^{\mu_{X2j}} \hat{\lambda}_{2j} \leq U'_j. \quad (3.34)$$

Equation (3.34) may be further approximated using

$\hat{\lambda}_{1j}^{\mu_{X1j}} = 1 + \hat{\lambda}_{1j} \ln(\mu_{X1j}) + R_j$ and obtaining an explicit expression for μ_{X1j}/μ_{X2j} ; however, this result indicates that inferences drawn on the original parameters are equivalent to those drawn on the transformed parameters. Although this may in fact occur when $\hat{\lambda}_1 \approx \hat{\lambda}_2 \approx 1$, this situation will not generally arise.

Returning to equation (3.34), let

$$Q_j = \{\mu_{X1j}, \mu_{X2j} : L'_j \leq \hat{\lambda}_{2j}^{\mu_{X1j}} \hat{\lambda}_{1j} - \hat{\lambda}_{1j}^{\mu_{X2j}} \hat{\lambda}_{2j} \leq U'_j\}. \quad (3.34a)$$

The pairs $(\mu_{X1j}, \mu_{X2j}) \in Q$ thus represent combinations of untransformed means that satisfy the inequality in equation (3.34) and Q may be graphed in terms of μ_{X1j} and μ_{X2j} as in Figure 1a. If Q contains some pairs, (μ_{X1j}, μ_{X2j}) , such that $\mu_{X1j} - \mu_{X2j} = 0$, these values may in fact be obtained by projecting the intersections of L_j and U_j with $(\mu_{X1j} - \mu_{X2j} = 0)$ onto the μ_{X1j} and μ_{X2j} axes, respectively, to obtain the closed intervals $[A_1, B_1]$ and $[A_2, B_2]$ as in Figure 1b. At this stage of the procedure it is possible to state that for $\mu_{X1j} \in [A_1, B_1]$ and $\mu_{X2j} \in [A_2, B_2]$ such that $\mu_{X1j} - \mu_{X2j} = 0$, the differences between the untransformed means is not significant. From Figure 1b, for example, the statement may be made that for $5 \leq \mu_{X1j} \leq 7.9$ and $5 \leq \mu_{X2j} \leq 7.9$,

$$\mu_{X1j} - \mu_{X2j} = 0.$$

Additional information may be obtained by taking advantage of the independence of μ_{X1j} and μ_{X2j} . By forming $(1-\alpha)$ simultaneous confidence regions for μ_{Y1j} and μ_{Y2j} using equation (3.29) or (3.30), interval estimates for the transformed means may be obtained:

$$L_{1j} \leq \mu_{Y1j} \leq U_{1j} ; \quad L_{2j} \leq \mu_{Y2j} \leq U_{2j} .$$

Again using the first order approximation to $E(Y_i) = f_i(\mu_{X1})$, the above equations reduce to:

$$L_{1j} \leq \frac{\mu_{X1j} \hat{\lambda}_{1j-1}}{\hat{\lambda}_{1j}} \leq U_{1j} ; \quad L_{2j} \leq \frac{\mu_{X2j} \hat{\lambda}_{2j-1}}{\hat{\lambda}_{2j}} \leq U_{2j} ,$$

which in turn reduce to

$$(\hat{\lambda}_{1j} L_{1j} + 1)^{1/\hat{\lambda}_{1j}} \leq \mu_{X1j} \leq (\hat{\lambda}_{1j} U_{1j} + 1)^{1/\hat{\lambda}_{1j}} \quad (3.34b)$$

and

$$(\hat{\lambda}_{2j} L_{2j} + 1)^{1/\hat{\lambda}_{2j}} \leq \mu_{X2j} \leq (\hat{\lambda}_{2j} U_{2j} + 1)^{1/\hat{\lambda}_{2j}} . \quad (3.34c)$$

Since μ_{X1j} and μ_{X2j} are independent and all $\hat{\lambda}_{ij}$ and L_{ij} in the above expressions are known for a given transformation function, these estimates may be computed and graphed as in Figure 1c. If the simultaneous confidence region, G_j , defined by the intersection of the intervals $[A'_1, B'_1]$ and $[A'_2, B'_2]$ calculated by equations (3.34b) and (3.34c), respectively, contains $\mu_{X1j} - \mu_{X2j} = 0$, then $H_0: \mu_{X1j} - \mu_{X2j} = 0$ is not

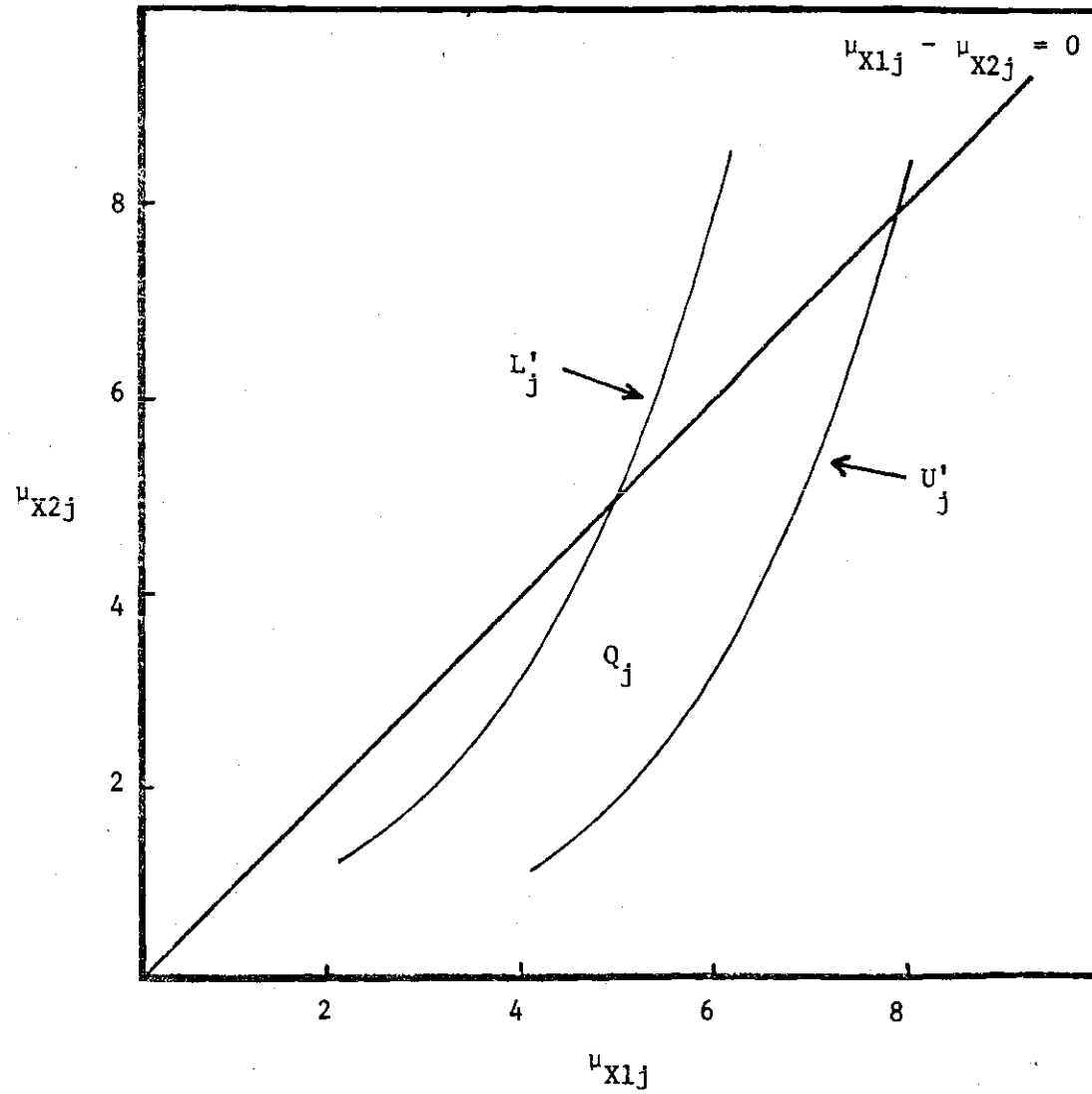


Figure 1a. Q Region for $(\mu_{x1j} - \mu_{x2j})$

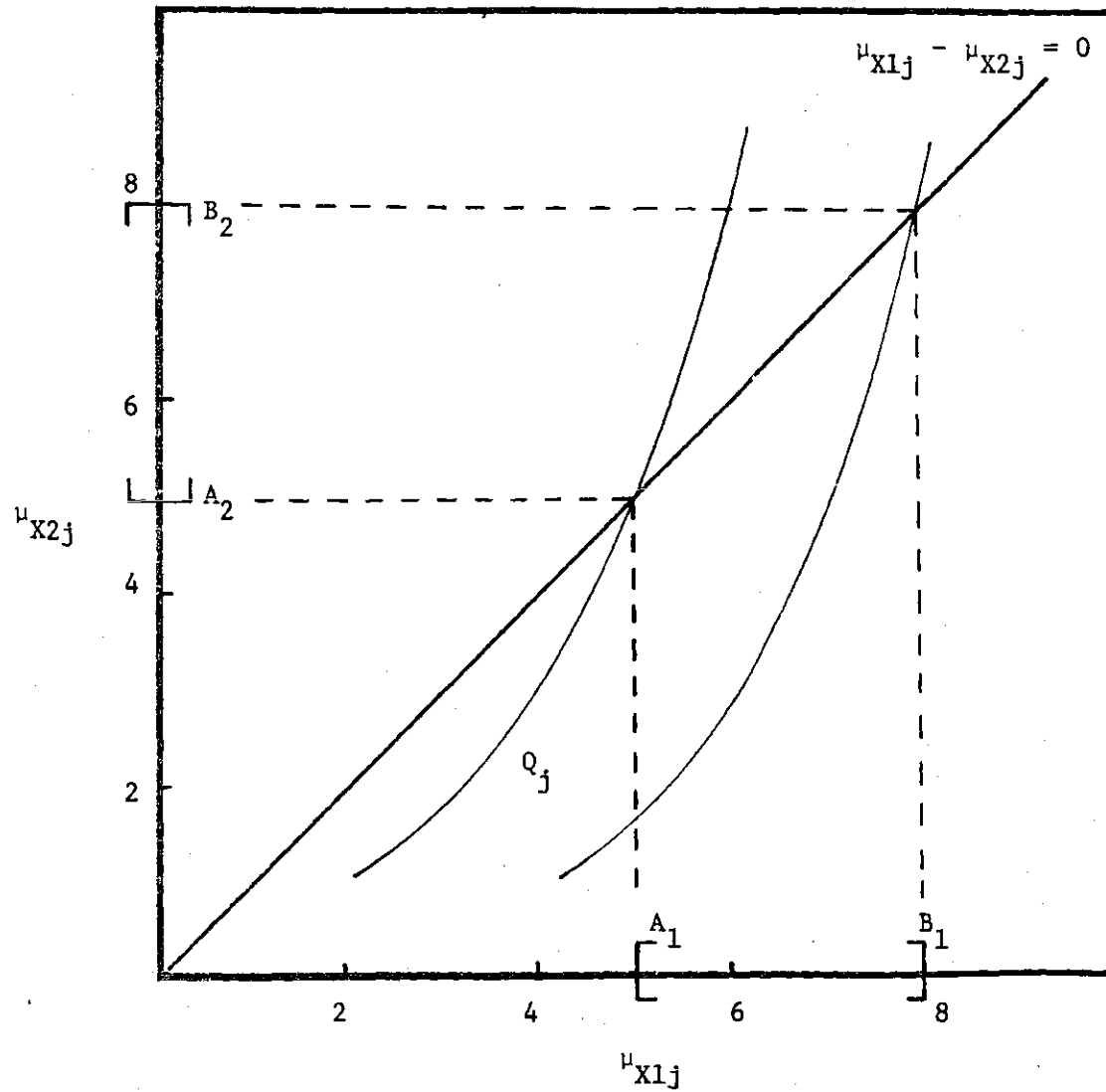


Figure 1b. Establishing Intervals for H_0 True

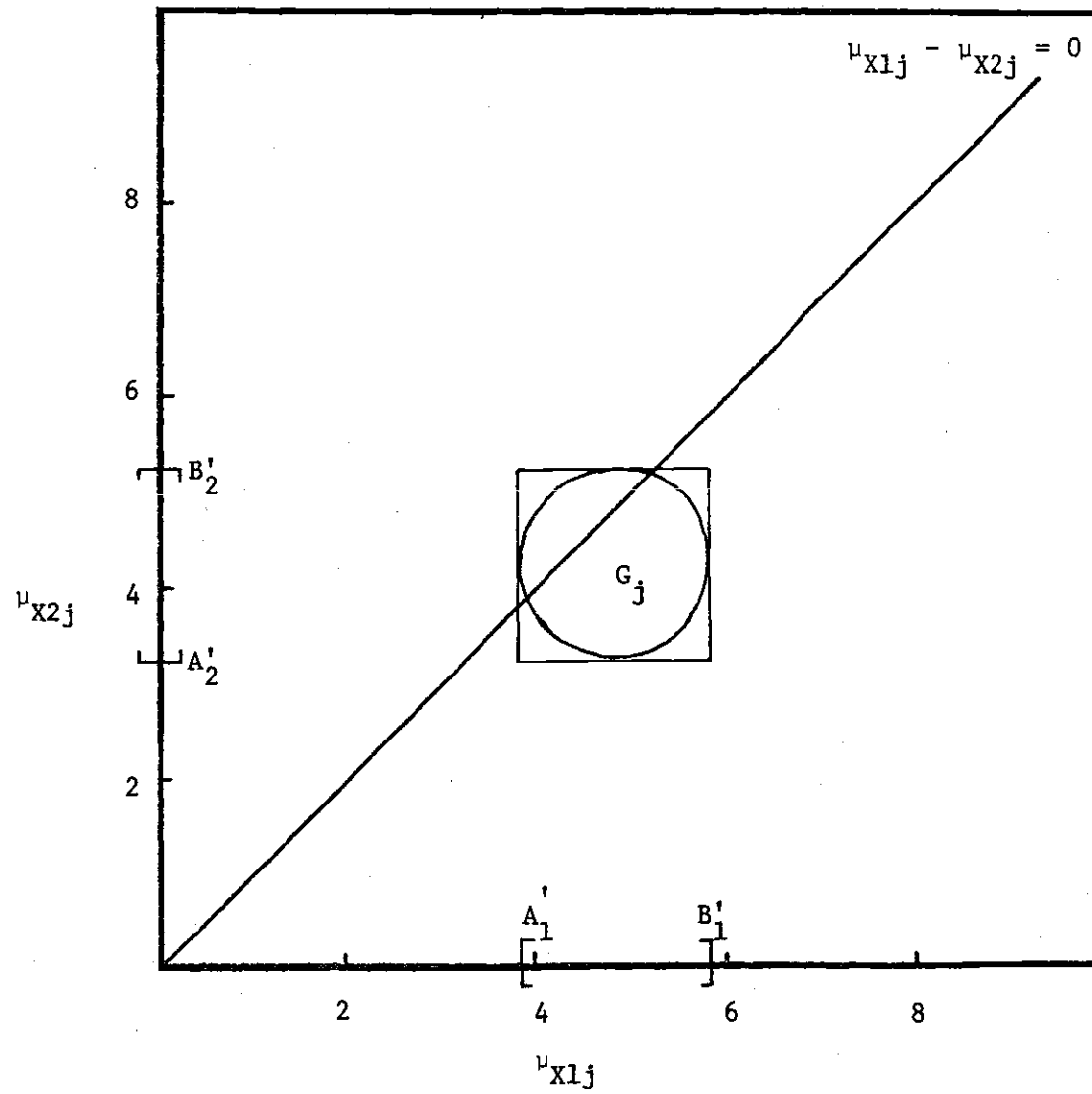


Figure 1c. G Region for $(\mu_{X1j} - \mu_{X2j})$

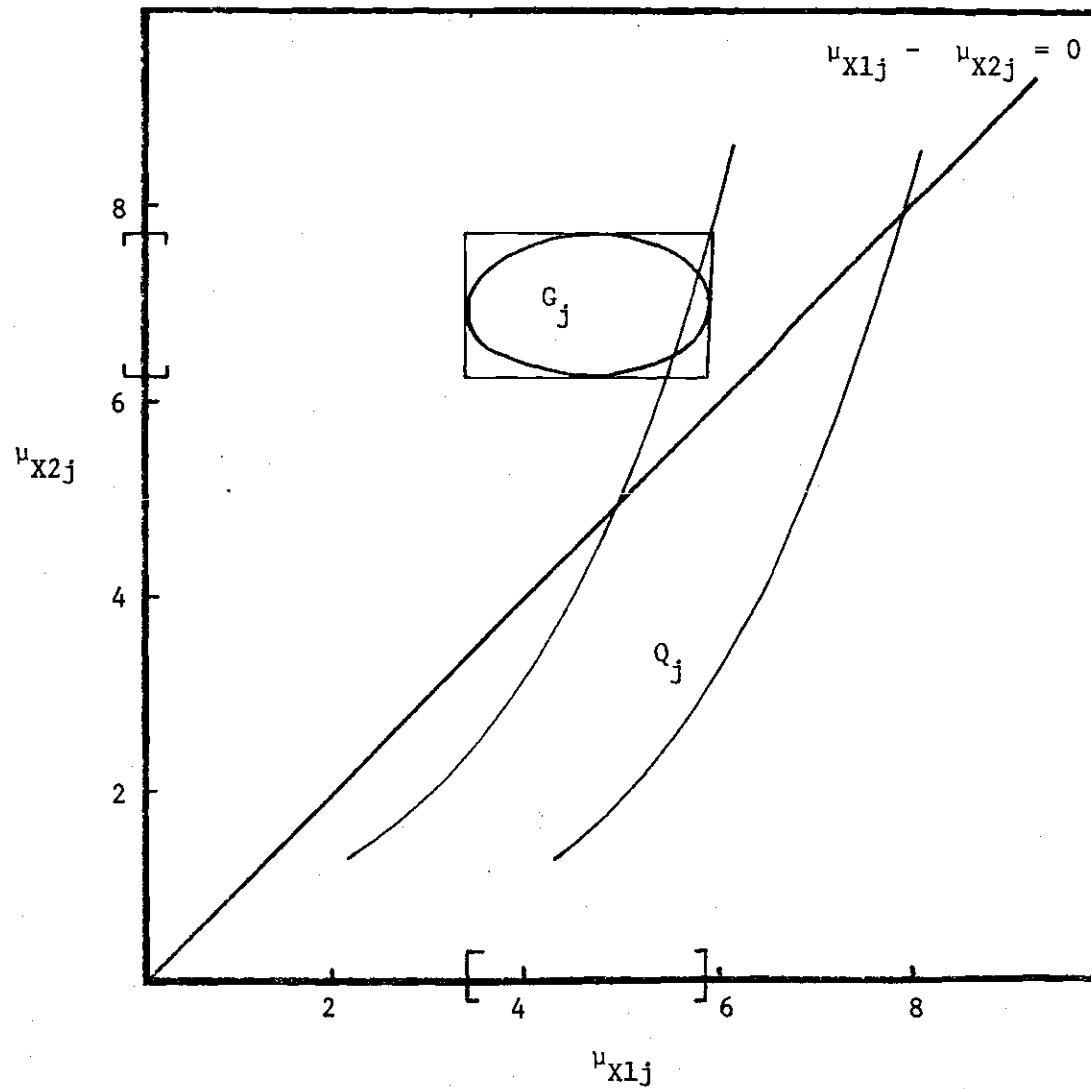


Figure 1d. Q and G Regions for $(\mu_{X1j} - \mu_{X2j})$

rejected with confidence approximately $(1-\alpha)$.

In examining the relationship between Q_j and G_j , numerous situations may arise. From Figure 1d, for example, G_j does not contain $\mu_{X1j} - \mu_{X2j} = 0$ while Q_j does indicate that for some pairs, (μ_{X1j}, μ_{X2j}) , $H_0: \mu_{X1j} - \mu_{X2j} = 0$ is not rejected. The conclusion in this situation is that although some pairs (μ_{X1j}, μ_{X2j}) have been identified for which $H_0: \mu_{X1j} - \mu_{X2j} = 0$ is not rejected, these pairs do not appear reasonable based on intervals established about μ_{X1j} and μ_{X2j} ; hence, $H_0: \mu_{X1j} - \mu_{X2j} = 0$ is rejected. Other specific situations will arise in the example problems examined in Chapter V and will be analyzed individually. Note should also be made of the fact that the Q region may have different shapes for different transformation parameters and it may happen that Q will not contain any portion of $(\mu_{X1j} - \mu_{X2j} = 0)$ for a particular i and j . Different Q regions will also be encountered in the example problems and will be examined in conjunction with their associated G regions on an individual basis.

A better approximation to Q may obviously be obtained by including second and higher order terms of $E(Y_{ij})$ from equation (3.32). The size and shape of Q would thus change depending on the magnitude of $\hat{\lambda}_{\sim i}$ and the variance of the untransformed variables, σ_{ij}^2 . Since

$$f_i''(\mu_{Xi}) = (\hat{\lambda}_{\sim i} - 1)\mu_{Xi}^{\hat{\lambda}_{\sim i}-2},$$

the second order term becomes

$$1/2(\hat{\lambda}_{\sim i}-1)\mu_{Xi}^{\hat{\lambda}_{\sim i}-2}\sigma_i^2.$$

For example problems considered in this research, $\hat{\lambda}_{i1}$ were relatively close to 1 and the sample variances, s_{j1}^2 , were fairly small so that inclusion of the second order term had little effect on Q. This in turn had no appreciable effect on inferences drawn on the original parameters based on this approximate method. If $\hat{\lambda}_{i1}$ are appreciably different from 1 and/or large sample variances, s_{j1}^2 , exist, then higher order terms should be included. Additional information will most likely be sacrificed even with the inclusion of higher order terms, since the true population variances, σ_{j1}^2 , are probably unknown and must be estimated by s_{j1}^2 .

The second case to be examined, $H_0: \mu_{Xi} - \mu_0 = 0$, is very similar to Case 1 in that an approximation procedure must be developed to preclude introduction of theoretical and computational difficulties mentioned earlier. A graphical solution is not necessary, however, for reasons to become immediately apparent. Recalling equation (3.33), the interval estimates of the transformed parameters may be approximated by

$$L_j \leq f_i(\mu_{Xij}) - \mu_{0j}^T \leq U_j \quad (3.35)$$

where the standard mean vector, μ_0 , has been transformed according to $\mu_0^T = f_i(\mu_0)$. Thus, equation (3.35) becomes

$$L_j \leq \frac{\mu_{Xij} \hat{\lambda}_{1j-1}}{\hat{\lambda}_{1j}} - \frac{\mu_{0j} \hat{\lambda}_{1j-1}}{\hat{\lambda}_{1j}} \leq U_j$$

which reduces to

$$\hat{\lambda}_{ij} L'_j \leq \mu_{Xij} - \mu_{0j} \leq \hat{\lambda}_{ij} U'_j, \text{ or}$$

$$L'_j \leq \mu_{Xij} - \mu_{0j} \leq U'_j.$$

Now, since $\mu_0^{\hat{\lambda}_{ij}}$ is a known constant,

$$L'_j + \mu_{0j}^{\hat{\lambda}_{ij}} \leq \mu_{Xij} \leq U'_j + \mu_{0j}^{\hat{\lambda}_{ij}}, \quad (3.36)$$

and

$$(L'_j + \mu_{0j}^{\hat{\lambda}_{ij}})^{1/\hat{\lambda}_{ij}} \leq \mu_{Xij} \leq (U'_j + \mu_{0j}^{\hat{\lambda}_{ij}})^{1/\hat{\lambda}_{ij}}. \quad (3.37)$$

If the interval estimate of μ_{Xij} from equation (3.37) contains μ_{0j} , then $H_0: \mu_{Xij} - \mu_{0j} = 0$ is true with confidence approximately $(1-\alpha)$. A slight computational difficulty arises, however, when either side of equation (3.36) is negative. This situation will be approached in the following manner:

1) If $L'_j + \mu_{0j} < 0$ and $U'_j + \mu_{0j} > 0$, set $L'_j + \mu_{0j} = 0$ and calculate the upper bound for μ_{Xij} using equation (3.37); then multiply equation (3.36) by -1 , calculate the lower bound, and multiply again by -1 to obtain the true lower bound.

2) If $L'_j + \mu_{0j} < 0$ and $U'_j + \mu_{0j} \leq 0$, multiply equation (3.36) by -1 , calculate the interval estimate of μ_{Xij} using equation (3.37), then multiply the result by -1 to obtain the true estimate of μ_{Xij} .

If any $\hat{\lambda}_{ij} = 0$, both cases discussed above become trivial. The transformation for this situation is $y_{ij} = \ln(x_{ij})$. For testing Case 1,

the interval estimates of the transformed means become

$$L_j \leq f_1(\mu_{X1j}) - f_2(\mu_{X2j}) \leq U_j,$$

$$L_j \leq \ln(\mu_{X1j}) - \ln(\mu_{X2j}) \leq U_j, \text{ or} \quad (3.38)$$

$$e^{L_j} \leq \frac{\mu_{X1j}}{\mu_{X2j}} \leq e^{U_j}$$

For testing $H_0: \mu_{X1j} - \mu_{0j} = 0$, equation (3.37) becomes

$$e^{L_j} \leq \frac{\mu_{X1j}}{\mu_{0j}} \leq e^{U_j} \quad (3.39)$$

Equations (3.38) and (3.39) obviously indicate that $1 \in [e^{L_j}, e^{U_j}]$ is equivalent to $0 \in [L_j, U_j]$; furthermore, the conclusions will always be identical.

The intricacies of these approximate methods will be discussed in Chapter V, so that continuation of the statistical review and general methodological development is now in order.

Principal Components

Principal component analysis was originally introduced by Pearson (31) and theoretically generalized by Hotelling (19, 20) and others (16) as a method for reducing the complexity of a p-variate system of data in order to facilitate further statistical analysis of relationships within the system. In general, principal components are merely those p linear combinations of the original variables that account for 100% of the variance in the system. With respect to the purpose of this paper, deter-

mination of relative contributions of the principal components will provide the basis for determining precise contributions of individual variables (MOE) to the systems of data describing operational effectiveness.

This technique is extremely useful since multivariate normality of the data is not required and, with one restriction to be discussed later, the principal components are unique. To facilitate interpretation of principal component analysis, presentation of the method will be geometrical in nature, as in (30); the interested reader is referred to (32) and (5) for precise theoretical considerations.

Consider the scatter diagram of N observations on each of the elements of a bivariate random variable, $\underline{X} = [X_1, X_2]$, as depicted in Figure 2.

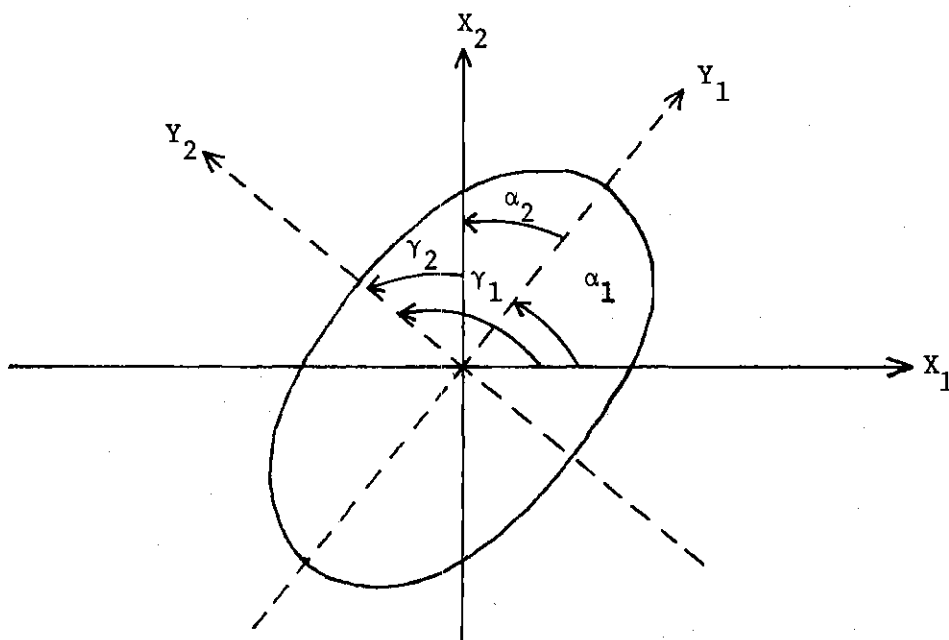


Figure 2. Scatter Diagram of Bivariate Observations

The swarm of observations is generally ellipsoidal in nature with major axis Y_1 and minor axis Y_2 . The angles α_1 and α_2 represent angles between Y_1 and each of the original response axes, X_1 and X_2 ; similarly, the angles γ_1 and γ_2 represent angles between Y_2 and each of the original response axes.

From analytic geometry, the orientation of Y_1 with respect to the original axes is completely determined by its direction cosines

$$a_{11} = \cos\alpha_1, \quad a_{21} = \cos\alpha_2;$$

the orientation of Y_2 may be similarly described by

$$a_{12} = \cos\gamma_1, \quad a_{22} = \cos\gamma_2,$$

where $a_{11}^2 + a_{21}^2 = 1$ and $a_{12}^2 + a_{22}^2 = 1$. Furthermore, the characteristic roots, $\ell = (\ell_1, \ell_2)$, of the sample covariance matrix, S , of the original data are known to be solutions to the determinantal equations

$$|S - \ell I|_{x_i} = 0,$$

where $x_i = (x_{i1}, x_{i2})$, $i = 1, 2$, are the characteristic vectors associated with each characteristic root, ℓ_i , of S (32).

If the major axis, Y_1 , is defined to be that axis that passes through the direction of maximum variance in the swarm, then Morrison (30) shows that solution of the expression for this variance

$$\frac{1}{N-1} \sum_{i=1}^N y_{i1}^2 = \frac{1}{N-1} \sum_{i=1}^N \left(\sum_{j=1}^2 a_{j1} (x_{ij} - \bar{x}_j)^2 \right) \quad (3.40)$$

will yield values for \underline{a}_1 which correspond to the elements of the characteristic vector associated with the largest characteristic root of \underline{S} . Solving an expression similar to equation (3.40) that maximizes the variance about the minor axis, Y_2 , will yield the characteristic vector, \underline{a}_2 , associated with the second largest characteristic root of \underline{S} . If ℓ_1 and ℓ_2 are distinct, then \underline{a}_1 and \underline{a}_2 are unique. The principal components Y_i , of the original set of data are then:

$$Y_1 = a_{11}X_1 + a_{21}X_2$$

$$Y_2 = a_{12}X_1 + a_{22}X_2,$$

or merely linear combinations of the original responses. In general, the i th principal component may then be defined as that linear combination of the original responses whose coefficients are the characteristic vectors associated with the i th greatest root, ℓ_i , of \underline{S} . The sample variance of the i th component has been shown to be ℓ_i (32), so that the total variance, V , of a p -dimensional system is:

$$V = \sum_{i=1}^P \ell_i \quad (3.41)$$

From equation (3.41) the contribution of the i th principal component, P_i , is obviously

$$P_i = \frac{\ell_i}{V} \quad (3.42)$$

Several iterative procedures are available for extracting the principal components from a given set of data (30) and will not be pre-

sented in this paper. Consideration will be given, however, to the question of whether the components should be extracted from the sample covariance matrix, \tilde{S} , or from the sample correlation matrix, \tilde{R} , since different results are obtained for each method. No exact guidelines for this situation have been established; however, consensus on the topic appears to be centered on the physical dimensionality of the data (32). If variables being examined represent similar measures (*e.g.*, metric measurements), then linear combinations of these measures are amenable to interpretation and computations based on \tilde{S} would seem appropriate. However, if different measures are used, then computations should probably be based on \tilde{R} to obtain dimensionless compounds of the original variates.

Contribution of Individual Variables

By reducing a p -dimensional system of data to linear combinations of the original variates that explain virtually all of the variance in the system, the original problem has been substantially simplified. The issue encountered is now one of providing an interpretation of principal components.

The majority of useful investigation in this area to date has centered around the identification of each principal component as a measurable variable that may be examined in further experimentation. Reduction of a large amount of data into a smaller set of identifiable variables that account for all, or nearly all, of the variance in the original data set was in fact the impetus for the development of this technique (19). Rao (34) and others (10) have provided additional procedures for identification of principal components that are essential-

ly extensions of the original methods, which involve extensive subjective evaluation.

With respect to the objective of this research, then, principal components would have to be identified as measurable MOE contributing to the operational effectiveness of a force. In certain circumstances, this may prove possible; however, since the MOE originally considered are assumed to completely describe operational effectiveness, further identification may be unjustified and attempts to do so may prove fruitless. Additional consideration, then, must be given to the mathematical characteristics of the principal components in an effort to determine the contribution of the original MOE to the operational effectiveness of a force.

The procedure developed herein to determine this contribution is based on the simple postulate that the contribution of a component of a system is merely the ratio of the corresponding weight of that element to the total weight of the system. Scaled percentages of the weights of all components of a system then account for 100% of the weight of the system. Two characteristics of principal components facilitate the exploitation of this postulate and will now be discussed.

Many interpretations of principal components, *e.g.*, (32), view the coefficients of the variates in each component as the "weight" of that variate in the component. This is similar to interpretations of regression coefficients as "weights" of terms in a regression equation. In light of this, the weight of a system in which all of its variance is accounted for in the first k principal components is merely the sum of all coefficients, a_{ij} in the k components. There is a drawback to

this particular interpretation, however, in that a large number of $a_{ij} < 0$ will result in a negative weight for the system. Account must also be given to the contribution of each principal component to the variance of the system.

This situation may be alleviated, however, by recalling that the principal components are orthogonal, *i.e.*, $\tilde{a}'\tilde{a} = 1$, which suggests an interest in a_{ij}^2 rather than a_{ij} . Also, the contribution of the variance of each principal component, P_i , to the variance of the total system may be found using equation (3.42). The total weight of the system may now be calculated as:

$$W = [P_1, P_2, \dots, P_k] \begin{bmatrix} a_{11}^2 & a_{21}^2 & \dots & a_{k1}^2 \\ a_{12}^2 & a_{22}^2 & \dots & a_{k2}^2 \\ \vdots & \vdots & \dots & \vdots \\ a_{1k}^2 & a_{2k}^2 & \dots & a_{kk}^2 \end{bmatrix}, \quad (3.43)$$

where the first k principal components account for 100% of the variance in the system. Since $\sum_{i=1}^k P_i = 1$ and $\tilde{a}'\tilde{a} = 1$, the total weight of any system is always 1. The contribution of any original variate, say X_i , to the total system is then

$$W_i = \frac{\sum_{j=1}^k P_j a_{ij}^2}{W}, \quad (3.44)$$

or simply

$$W_i = \sum_{j=1}^k P_j a_{ij}^2. \quad (3.44a)$$

In actuality, then, this procedure merely translates the percentage contribution of principal components into percentage contribution of each variable to the total variance of the system of data being examined.

Multivariate Random Vector Generation

Since no actual data were available from stochastic simulation of the example problems to be presented in Chapter V, observations had to be generated in order to demonstrate the methodology. Although observations needed are thought to follow a distribution other than $N(\mu, \Sigma)$, the author was forced to generate observations from $N(\mu, \Sigma)$ due to availability of requisite computer programs. This limitation proved advantageous, however, when assessing the merits of tests for multivariate normality.

The generation was based on the fact that $\tilde{X} \sim N(\mu_X, \Sigma_X)$ if and only if

$$\tilde{X} = \tilde{C}\tilde{Z} + \mu_X, \quad (3.45)$$

where $\tilde{Z} \sim N(0, I)$ and \tilde{C} is a unique lower triangular matrix such that $\Sigma_X = \tilde{C}\tilde{C}'$ (5). The generation of \tilde{Z} was accomplished by using the following equations from Fishman (14):

$$x_1 = (-2 \ln U_1)^{1/2} \cos (2\pi U_j) \quad (3.46)$$

$$x_j = (-2 \ln U_1)^{1/2} \sin (2\pi U_j), \quad (3.46a)$$

where U_1 and U_j are independent deviates from $U(0, 1)$. The \tilde{C} matrix may

be computed by the square root method attributed to Scheuer and Stoller

(37). The generation of \tilde{X} from $N(\mu_{\tilde{X}}, \Sigma_{\tilde{X}})$ is thus accomplished by:

1. Generation of independent variates from $U(0, 1)$;
2. Generation of \tilde{Z} from (3.46) and (3.46a);
3. Generation of \tilde{C} ; and
4. Computation of \tilde{X} using (3.45).

CHAPTER IV

DEVELOPMENT OF THE SPECIFIC METHODOLOGY

Introduction

The purpose of this chapter is to develop the specific methodology for accomplishing the primary objective stated in Chapter I based on a synthesis of statistical methods reviewed in Chapter III. The methodology essentially consists of two portions:

1. Ascertaining the presence of statistically significant differences between two sets of MOE describing the operational effectiveness of a military force, and
2. Determining the contribution of individual MOE to a force's operational effectiveness and the contribution of differences in MOE for two systems to differences in operational effectiveness.

As mentioned previously, the results of the first portion may preclude the necessity for implementation of the second; similarly, the second portion may be executed without implementation of the first, *i.e.*, the two sections supplement one another, but may be executed independently.

Determination of Significant Differences

Given two sets of data, X_1 and X_2 , collected from independently conducted stochastic simulations of hypothesized combat, consideration must first be given to the respective underlying distributions of the random variables (MOE) being examined. If a statement regarding the multivariate normality of each set, X_1 , may be made, then further anal-

ysis may proceed. The multivariate W-statistic developed by Malkovich and Afifi, equation (3.9), and the likelihood statistic developed by Andrews, *et. al.*, equation (3.12), actually provide two methods from which statements concerning multivariate normality of a vector random variable may be made.

The procedure adopted will be to first compute the W-statistic for each data set and ascertain the appropriate confidence level associated with these statistics from Table 21, Appendix A. If either or both data sets satisfy pre-established criterion for assuming multivariate normality, then the raw data may be used in further analyses. However, if either or both sets do not satisfy the criterion, then joint multivariate normality must be induced in the set that violates the criterion so that hypothesis tests based on $X_i \sim N(\mu_{Xi}, \Sigma_{Xi})$ may be used.

For this purpose, the non-linear transformation $Y_i = f_i(X_i)$, equation (3.10), will be used to determine the set of transformation parameters, $\hat{\lambda}_i$, that maximizes $L_{\max}(\lambda_i)$. The confidence level associated with $\hat{\lambda}_i$ will then be computed using equation (3.12). If both X_1 and X_2 fail the initial test based on the W-statistic, then the transformation will be applied separately to each set of variables to obtain two sets of parameters, $\hat{\lambda}_1$ and $\hat{\lambda}_2$. In this way, confidence levels may be obtained for each set of parameters which would be higher than a composite $\hat{\lambda}$ obtained by maximizing $L_{\max}(\lambda_1)$ and $L_{\max}(\lambda_2)$ simultaneously. As a means of checking the validity of transformation parameters obtained through this maximization procedure, the W-statistic will again be calculated for those data sets that have been transformed to determine a new confidence level based on W which may be compared to that associated with

$\hat{\lambda}_1$.

Once multivariate normality has been verified in and/or induced on each set of variables, hypothesis testing may begin. By first testing for equality of covariance matrices

$$H_0: \Sigma_{X1} = \Sigma_{X2}$$

$$\text{against } H_1: \Sigma_{X1} \neq \Sigma_{X2}$$

the appropriate test for the differences in mean vectors may be selected.

If each X_i is of dimension $p \leq 4$ or 5 and $N_i < 20$, then equations (3.14) and (3.15) will be used to compute MC^{-1} which will be compared to $\chi^2_{1-\alpha; p}$. If each X_i is of the dimension $p \geq 5$ or $N_i > 20$, then equations (3.14) and (3.18) will be used to compute Mb^{-1} which will be compared to

$F_{1-\alpha; f_1, f_2}$ as proposed by Box.

If $H_0: \Sigma_{X1} = \Sigma_{X2}$ is not rejected with confidence $(1-\alpha)$, then testing

$$H_0: \mu_{X1} = \mu_{X2}$$

$$\text{against } H_1: \mu_{X1} \neq \mu_{X2}$$

will proceed in the usual manner by comparing T^2 against T_0^2 using equations (3.20) and (3.21). If $H_0: \Sigma_{X1} = \Sigma_{X2}$ is rejected, then Bennett's procedure for solving the multivariate Behrens-Fisher problem will be used to test the differences of the means. In this situation, equations (3.27) and (3.28) will be used to calculate T^2 and T_0^2 to provide the

necessary comparative statistics.

If a standard vector of means is available, then testing

$$H_0: \mu_{\underline{X}_i} = \mu_0$$

$$\text{against } H_1: \mu_{\underline{X}_i} \neq \mu_0$$

would be appropriate in an effort to obtain more information about potential differences between \underline{X}_1 and \underline{X}_2 . Statistics for testing these hypotheses will be calculated using equations (3.23) and (3.24) for comparing T^2 and T_0^2 .

If both data sets satisfy the initial criterion for the test of multivariate normality based on the first W-statistics computed, then inferences drawn on the above hypothesis tests about the means are valid. For those hypotheses that are rejected, simultaneous confidence intervals will be established about $(\mu_{\underline{X}_1} - \mu_{\underline{X}_2})$ and $(\mu_{\underline{X}_i} - \mu_0)$ using equations (3.29) and (3.30), respectively, to determine which elements of \underline{X}_1 and \underline{X}_2 are in fact contributing to the rejection of the associated null hypotheses. Inferences drawn from these interval estimates are also valid if neither data set had to be subjected to a transformation to insure multivariate normality.

However, if either or both data sets required transformation prior to proceeding with hypothesis tests on the means, inferences drawn from these tests may or may not be equivalent to inferences drawn with respect to the parameters of the original variables. In this case the approximate procedures developed in Chapter III will be utilized to equate inferences about parameters of transformed variables to inferences about

parameters of the original variables. Regardless of whether or not hypotheses on the means of transformed variables are rejected, simultaneous confidence intervals will be established using equations (3.29) and (3.30). This will protect against failure to recognize differences between parameters of the original variables when tests about differences of the parameters of the transformed variables are not rejected.

Computationally, for intervals constructed about $(\mu_{Y1} - \mu_0)$, equation (3.37) may be used directly to calculate intervals about $(\mu_{X1} - \mu_0)$, after checking to insure that $L'_j + \mu_{0j}^{\lambda_{ij}} > 0$ and $U'_j + \mu_{0j}^{\lambda_{ij}} > 0$ for $j = 1, 2, \dots, p$. For intervals constructed about $(\mu_{Y1} - \mu_{Y2})$, the graphical procedure introduced in Chapter III will be appropriate. To avoid confusion, a separate graph should be prepared for each $(\mu_{Y1j} - \mu_{Y2j})$, $j = 1, 2, \dots, p$. Using these approximate methods will thus allow inferences to be made with respect to parameters of the original variables. Analysis of these results is fairly straightforward and follows directly from the discussion in Chapter III. Possible statistical differences between MOE being examined will now have been identified so that determination of individual MOE contributions may be examined. The reader is reminded that continuing at this point may not be appropriate if no differences between MOE have been substantiated.

Contribution of Individual MOE

The second portion of the methodology follows directly from the discussion of principal component analysis in Chapter III. After determination of whether principal components should be extracted from sample covariance matrices, S_1 , or from sample correlation matrices, R_1 , computations may proceed. Since multivariate normality is not required for

computation of principal components, and no inferences about the latent roots or eigenvectors are to be made, computations will be based on maximum likelihood estimates of $S_{\sim 1}$ calculated from the original data using equation (3.6).

In order to obtain as much information as possible about individual MOE contributions, principal components will be extracted from both $X_{\sim 1}$ and $X_{\sim 2}$, in addition to $(X_{\sim 1} - X_{\sim 2})$. Once this has been accomplished, contributions of individual MOE to the variance of each $X_{\sim 1}$ and contributions of differences in MOE to $(X_{\sim 1} - X_{\sim 2})$ will be calculated using equation (3.44a). In this way, the contribution of both sets of MOE to the operational effectiveness of a force may be compared directly, in addition to the establishment of the contribution of differences in MOE to differences in operational effectiveness. Reminder is made of the fact that analysis of the results of this section of the methodology should encompass an investigation of the coefficients of the principal components in conjunction with percentage contributions of individual MOE and differences in MOE.

Methodological Input

Prior to implementing the methodology described in the previous two sections, several decisions must be made at an appropriate level with respect to the desired exactness of the procedure.

1. An adequate sample size must be determined so that sufficient statistical information may be obtained from the results of each replication of the stochastic simulation.

2. Appropriate confidence levels must be established for testing multivariate normality, equality of covariance matrices, and differences

of mean vectors. These confidence levels may be different for reasons mentioned previously.

3. If possible, a standard vector of means should be established for the forces being considered so that as much information as possible may be obtained from the first portion of the methodology.

4. Determination must be made with respect to extraction of principal components, *i. e.* from the sample covariance matrices or the sample correlation matrices.

Although none of the aforementioned requirements affects the methodological procedure, consideration must be given so that the best information available may be obtained from the data. No precise consideration will be given to these requirements; rather, these decisions will be assumed to have been made prior to implementation of the methodology, which will now be demonstrated through two example problems.

The entire methodology, except graphical computations, has been programmed in FORTRAN IV by the author for use on a CDC 7000 computer and is portrayed in Figure 3. The complete program listing is in Appendix 2, except for those portions that utilize prepared subroutines or function statements obtained from (43) and (46). That portion of the methodology that concerns reconstruction of confidence intervals in terms of the parameters of the original variables was performed on a desk calculator when necessary.

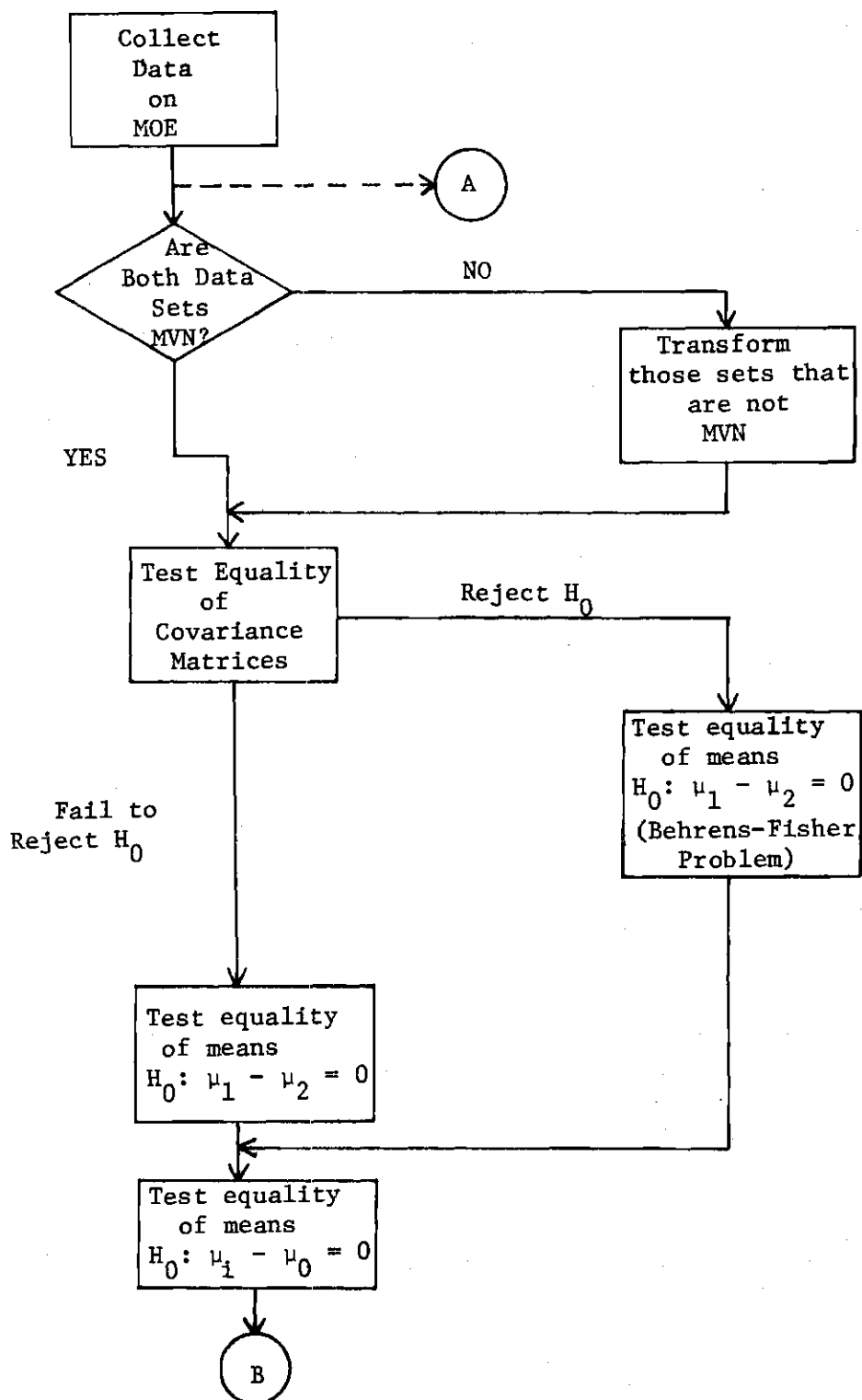


Figure 3. Flow Diagram of the Methodology

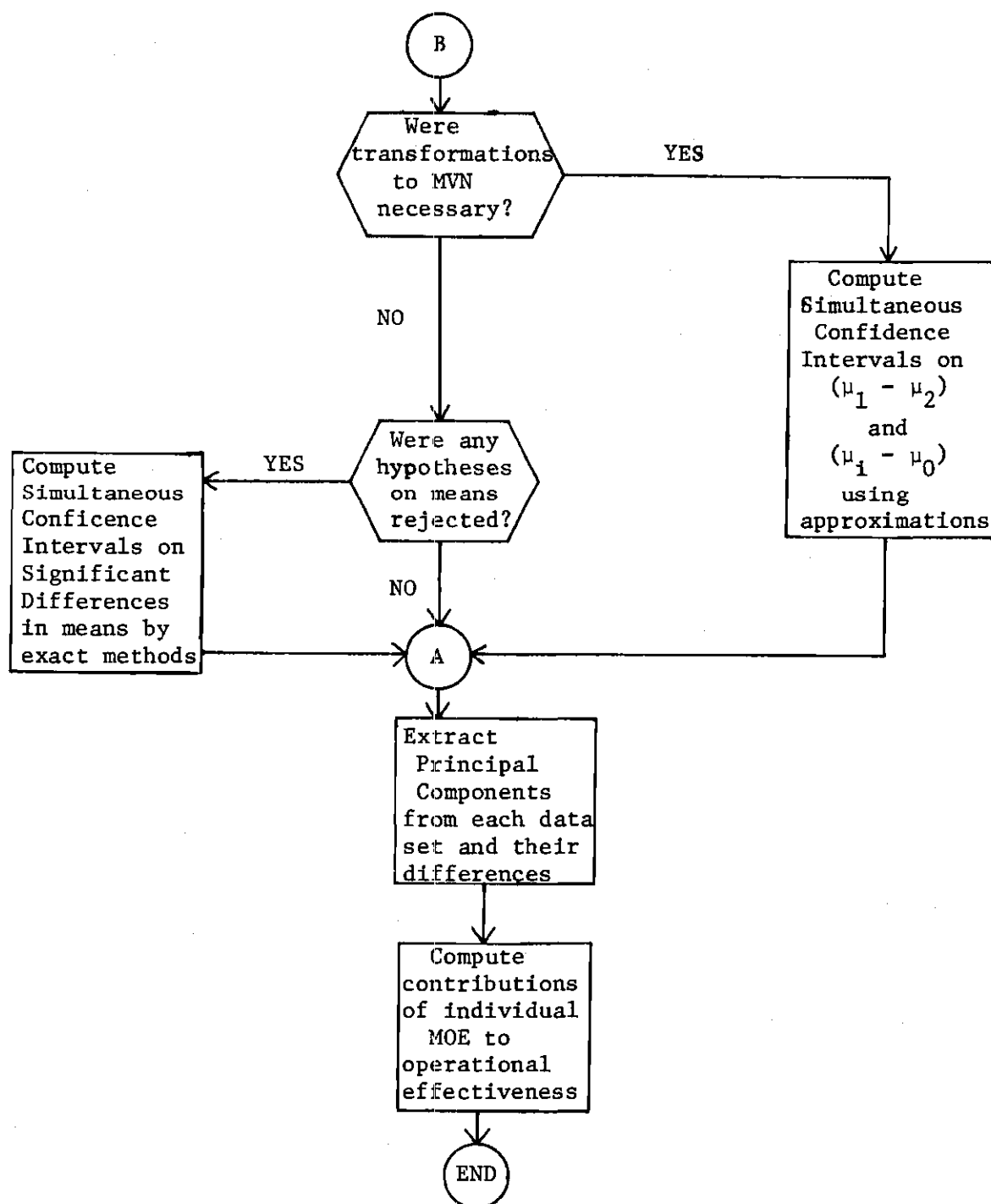


Figure 3 (continued). Flow Diagram of the Methodology

CHAPTER V

DEMONSTRATION OF THE METHODOLOGY

Introduction

In this chapter the methodology summarized in Chapter IV will be demonstrated by the examination of two practical, albeit hypothetical, problems. The primary purpose in so doing, of course, will be to demonstrate the methodology; the secondary purpose will be to examine the two definitions of operational mobility presented in Chapter II.

Background of the Problems

Consider a situation in which the U. S. Army has validated a requirement for a new combat fighting vehicle (CFV-2) designed to replace the current vehicle (CFV-1) employed in mechanized infantry battalions. To obtain as much information as possible on the effectiveness of CFV-2, the Commander, TRADOC, has tasked an Army agency to conduct a COEA. This analysis is to be conducted in conjunction with phase one of operational testing (OT-I) in order to establish comparative measures between the operational effectiveness of a mechanized infantry battalion using CFV-1 and the effectiveness of the same force using CFV-2.

Since CFV-2 has been designed to improve the mobility of ground forces during combat, the commander, TRADOC has directed that measures of each force's operational mobility be included in the operational effectiveness analysis portion of the COEA.

Data for the analysis will be collected from twenty replications

of a stochastic simulation model designed for examination of a mechanized infantry battalion in hypothesized combat roles. Three battle scenarios will be examined in two combat environments. For example, twenty replications of the experiment will be obtained for a mechanized infantry battalion in the attack using CFV-1 in Environment-I.

With respect to the issue of operational mobility of the forces being examined, two sets of MOE have been approved to represent the operational effectiveness of the force. Set 1 will consist of four MOE, designated MOE_1 through MOE_4 , in which MOE_4 will be the measure of the force's momentum. Set 2 will consist of eight MOE, designated MOE_1 through MOE_8 . The first three MOE for each set are the same, while MOE_4 through MOE_8 of Set 2 represent a subset of critical operational mobility performance characteristics common to CFV-1 and CFV-2. Performance characteristics of other vehicles within the force will be held constant throughout the experiment, solely for purposes of demonstrating the methodology.

For purposes of data analysis, confidence levels for hypothesis tests, excluding the tests for multivariate normality, will be 95%. For the tests of multivariate normality, any positive indication of normality will be accepted. In addition, principal components will be extracted from the sample covariance matrices of the respective data sets since MOE will be in terms of dimensionless LER.* Also, standard vectors of means for each set of MOE have been approved by the Commander, TRADOC:

*If MOE were defined some other way, say in terms of force differences, then principal components should probably be extracted from sample correlation matrices.

	MOE ₁	MOE ₂	MOE ₃	MOE ₄
SET 1:	[3.2,	3.8,	2.4,	2.0]

	MOE ₁	MOE ₂	MOE ₃	MOE ₄	MOE ₅	MOE ₆	MOE ₇	MOE ₈
SET 2:	[5.2,	5.8,	4.5,	3.3,	3.1,	2.3,	3.0,	3.9]

After collection of all data, the statistical analysis may proceed. For demonstration purposes only, the following analyses will examine only hypothesized results of one scenario-environment combination for forces using CFV-1 and CFV-2 for each set of MOE. Analyses of other combinations would proceed in exactly the same manner.

Example Problem I

This analysis will be concerned with an examination of hypothesized simulation results of the effectiveness of a mechanized infantry force using CFV-1 and CFV-2 in terms of MOE Set 1, in which the operational mobility of the force is measured in terms of relative momentum. Numerical results have been obtained from the computer program listing in Appendix B. In the following analysis, the subscripts 1 and 2 correspond to data pertaining to CFV-1 and CFV-2, respectively.

Section I

In the initial assessment of multivariate normality for the MOE obtained for each vehicle, the associated W-statistic must be computed from equation (3.9):

$$W_i = \frac{\left[\sum_{j=1}^n a_{ij} U_j \right]^2}{(X_i - \bar{x})' A^{-1} (X_i - \bar{x})}$$

The order statistics, U_j , calculated from equation (3.8), are displayed in Table 1. Obtaining the normalized coefficients, $\{a_{n-1+1}\}$, from Table 20, Appendix A, and computing W_1 gives

$$W_1 = .9317$$

$$W_2 = .8378$$

Table 1. Order Statistics for X_1 and X_2

<u>CFV-1</u>		<u>CFV-2</u>	
-.2090	.0249	-.2333	-.0004
-.2052	.0330	-.2073	.0035
-.1673	.0343	-.1597	.0065
-.1592	.0453	-.1596	.0215
-.1207	.0836	-.1192	.0446
-.1120	.0862	-.1187	.0645
-.0703	.1126	-.0394	.0722
-.0391	.1246	-.0157	.0781
-.0087	.1358	-.0096	.2239
.0163	.3951	-.0087	.5569

Note should be taken of the fact that the coefficients are symmetric, *i.e.*, $-a_{n-i+1} = a_i$. Referring to Table 21, Appendix A to determine the levels of confidence associated with each W_i reveals that $W_1 \approx W_{20}$ at the 20% level, while $W_2 < W_{20}$ at the 1% level. Although the interpretation is favorable for X_1 based on W_1 , and unfavorable for W_2 , the assumption will be made at this point that neither X_1 nor X_2 is multivariate normal so that all facets of the methodology may be demonstrated.

Applying the transformation

$$Y_i = \begin{cases} \frac{X_i^{\lambda_i} - 1}{\lambda_i} & , \lambda_i \neq 0 \\ \ln(X_i) & , \lambda_i = 0 \end{cases}$$

to each $X_{\sim 1}$ separately, maximizing

$$L_{\max}(\lambda_{\sim 1}) = -\frac{n_{\sim 1}}{2} \ln |S_{\sim 1}| + \left(\sum_{j=1}^4 (\lambda_j - 1) \sum_{k=1}^N \ln x_{ijk} \right)$$

now becomes necessary in order to induce the multivariate normality of $X_{\sim 1}$ and $X_{\sim 2}$. The algorithm adopted for this procedure is a cyclic coordinate method due to Bazaara (6). The transformation parameters obtained are

$$\hat{\lambda}_{\sim 1} = [.92, .95, 1.10, 1.58]$$

and

$$\hat{\lambda}_{\sim 2} = [.98, .98, .59, .59] ,$$

giving

$$L_{\max}(\hat{\lambda}_{\sim 1}) = 51.7963, \quad L_{\max}(\hat{\lambda}_{\sim 2}) = 45.3018$$

Computing $L_{\max}(1)_{\sim 1} = 50.0283$ and $L_{\max}(1)_{\sim 2} = 42.7555$ yields

$$2 [L_{\max}(\hat{\lambda}_{\sim 1}) - L_{\max}(1)_{\sim 1}] = 3.5359 ,$$

$$\text{and } 2 [L_{\max}(\hat{\lambda}_{\sim 2}) - L_{\max}(1)_{\sim 2}] = 5.0926;$$

calculating $(1-\alpha)$ for X_4^2 , gives the confidence level associated with $\hat{\lambda}_{\sim 1}$ as 47% and that associated with $\hat{\lambda}_{\sim 2}$ as 28%.

Re-calculating W_1 and W_2 for the transformed variables gives

$$W_1 = .9456$$

$$W_2 = .8991 .$$

Returning again to tables of the W distribution reveals that $W_1 \approx W_{20}$ at approximately the 45% level, while $W_2 \approx W_{20}$ at approximately the 5% level. The confidence levels associated with $X_{\sim 1}$ are observed to be approximately the same and the W-statistic has more than doubled after the transformation. Although the confidence levels for $X_{\sim 2}$ are not in the

same vicinity, the W-statistic for the transformed variables now indicates multivariate normality with 5% confidence, which is substantially better than no indication prior to transformation. Low levels of confidence associated with assuming $\underline{Y}_2 \sim N(\underline{\mu}_{Y2}, \underline{\Sigma}_{Y2})$ may be attributed to the sensitivity of the W statistic as mentioned previously. The efficiency of the particular non-linear maximization procedure used also affects the associated confidence levels and will be discussed in Chapter VI. However, since positive indications of multivariate normality have been achieved, the assumptions that $\underline{Y}_1 \sim N(\underline{\mu}_{Y1}, \underline{\Sigma}_{Y1})$ and $\underline{Y}_2 \sim N(\underline{\mu}_{Y2}, \underline{\Sigma}_{Y2})$ will be made and hypothesis testing on the transformed variables will be given.

In testing equality of covariance matrices

$$H_0: \underline{\Sigma}_{Y1} = \underline{\Sigma}_{Y2}$$

the statistic MC^{-1} will be compared to $\chi^2_{.05, 10}$ since $\underline{\Sigma}_{Y1}$ and $\underline{\Sigma}_{Y2}$ are of dimension $p = 4$ and $N_1 = N_2 = 20$. Using equations (3.14) and (3.15) gives

$$M = \left(\sum_{i=1}^2 n_i \right) \ln |\underline{S}_p| - \sum_{i=1}^2 n_i \ln |\underline{S}_i| = 57.3539 ,$$

and

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left(\sum_{i=1}^k \frac{1}{n_i} - \frac{1}{\sum_{i=1}^k n_i} \right) = .8868 ,$$

so that $MC^{-1} = 50.8638$, which is greater than $\chi^2_{.05, 10} = 18.307$. The null hypothesis is thus rejected at the 95% level of confidence and the

Behrens-Fisher problem has been encountered for testing $H_0: \mu_{Y1} - \mu_{Y2} = 0$.

The statistics necessary for testing $H_0: \mu_{Y1} - \mu_{Y2} = 0$ are

$$T^2 = N\bar{y}' S^{-1} \bar{y}$$

and

$$T_0^2 = \frac{p(N-1)}{N-p} F_{.05; 4, 16}$$

from equations (3.27) and (3.28). These calculations give $T^2 = 255.4379$ and $T_0^2 = 14.2829$, so the null hypothesis is rejected at the 95% level of confidence. The conclusion is made that some, or all, of the transformed means of the MOE for CFV-1 and CFV-2 are different.

Tests will now be made in an effort to determine if either set of transformed means of the MOE for CFV-1 and CFV-2 are different from the transformed standard mean vector. In testing $H_0: \mu_{Y1} - \mu_0^T = 0$, statistics necessary are

$$T^2 = N(\bar{y}_1 - \mu_0^T)' S_1^{-1} (\bar{y}_1 - \mu_0^T)$$

and

$$T_0^2 = \frac{p(N-1)}{N-p} F_{.05; 4, 16}$$

from equations (3.23) and (3.24), respectively. Calculations yield $T^2 = 1294.1079$, $T_0^2 = 14.2829$ so that the null hypothesis is rejected at the 95% level of confidence. Similar calculations for testing $H_0: \mu_{Y2} - \mu_0^T = 0$ give $T^2 = 669.0139$ and $T_0^2 = 14.2829$, so that this null

hypothesis is also rejected. The results of these three tests indicate that not only are the means of each set of transformed MOE different from the transformed standard mean vector, but they are also different from each other.

Rejection of these three hypotheses dictates the necessity for construction of simultaneous confidence intervals about the detected differences in the transformed means. This procedure will indicate which specific transformed means are in fact different and will provide the basis for translating these inferences into inferences on the means of the original MOE.

For the differences $(\mu_{Y1} - \mu_0^T)$ and $(\mu_{Y2} - \mu_0^T)$, equation (3.30) gives the results displayed in Table 2.

Table 2. Confidence Intervals for $(\mu_{Yij} - \mu_{0j}^T)$

<u>CFV-1</u>	<u>CFV-2</u>
$-1.0234 \leq \mu_{Y11} - \mu_{01}^T \leq .5563$	$-1.5328 \leq \mu_{Y21} - \mu_{01}^T \leq 1.0657$
$-1.0693 \leq \mu_{Y12} - \mu_{02}^T \leq 1.0111$	$-1.4365 \leq \mu_{Y22} - \mu_{02}^T \leq 1.3783$
$-1.0937 \leq \mu_{Y13} - \mu_{03}^T \leq -.1444$	$-1.0009 \leq \mu_{Y23} - \mu_{03}^T \leq -.2372$
$.0118 \leq \mu_{Y14} - \mu_{04}^T \leq 2.1259$	$.6030 \leq \mu_{Y24} - \mu_{04}^T \leq 1.5347$

Using equation (3.37),

$$(L'_j + \mu_{0j})^{\hat{\lambda}_{ij}} \frac{1}{\hat{\lambda}_{ij}} \leq \mu_{Xij} \leq (U'_j + \mu_{0j})^{\hat{\lambda}_{ij}} \frac{1}{\hat{\lambda}_{ij}}$$

to obtain intervals on the means of the original MOE gives the intervals displayed in Table 3.

Since $(\mu_{X13} - \mu_{03})$, $(\mu_{X14} - \mu_{04})$, $(\mu_{X23} - \mu_{03})$ and $(\mu_{X24} - \mu_{04})$

do not contain 0, the conclusion is made that the means of MOE_3 and MOE_4 for both CFV-1 and CFV-2 are different from the corresponding elements of the standard mean vector. Simultaneous confidence intervals must now be constructed about $(\mu_{Y1} - \mu_{Y2})$ to determine which means for CFV-1 and CFV-2 are different from each other.

Table 3. Confidence Intervals for $(\mu_{X1j} - \mu_0)$

<u>CFV-1</u>	<u>CFV-2</u>
$-1.1055 \leq \mu_{X11} - \mu_{01} \leq .6158$	$-1.5599 \leq \mu_{X21} - \mu_{01} \leq 1.0916$
$-1.1327 \leq \mu_{X12} - \mu_{02} \leq 1.0907$	$-1.4695 \leq \mu_{X22} - \mu_{02} \leq 1.4170$
$-1.0280 \leq \mu_{X13} - \mu_{03} \leq -.1330$	$-1.2505 \leq \mu_{X23} - \mu_{03} \leq -.3296$
$.0081 \leq \mu_{X14} - \mu_{04} \leq 1.2218$	$.8657 \leq \mu_{X24} - \mu_{04} \leq 2.4436$

Using equation (3.30) again, based on the calculations from the Behrens-Fisher problem, gives the confidence intervals displayed in Table 4.

Table 4. Confidence Intervals for $(\mu_{Y1j} - \mu_{Y2j})$

$-2.6116 \leq \mu_{Y11} - \mu_{Y21} \leq .5279$
$-2.5052 \leq \mu_{Y12} - \mu_{Y22} \leq .7205$
$-.6057 \leq \mu_{Y13} - \mu_{Y23} \leq .6010$
$.2200 \leq \mu_{Y14} - \mu_{Y24} \leq 2.5489$

The approximate graphical procedure developed in Chapter III will now be employed so that inferences may be drawn on the differences in the means of the original MOE. In order to graph the Q-region for each interval, equation (3.34a) was solved for μ_{X1j} in terms of μ_{X2j} giving two equations corresponding to the lower and upper limits of Q. These equations were

then solved for μ_{X1j} for integer values of μ_{X2j} between 0 and 10. The intervals, $[A_1', B_1']$ and $[A_2', B_2']$, necessary for obtaining the G_j regions, $j = 1, 2, 3, 4$, may be obtained by solving equations (3.34b) and (3.34c) directly. The results of these computations are portrayed in Figures 4-7.

In Figure 4, preliminary indications are that $H_0: \mu_{X11} - \mu_{X21} = 0$ should not be rejected since $\mu_{X11} - \mu_{X21} = 0$ is contained entirely in Q_1 . This is substantiated by the fact that $\mu_{X11} - \mu_{X21} = 0$ is also in G_1 , so the conclusion is made that $H_0: \mu_{X11} - \mu_{X21} = 0$ is not rejected with confidence approximately $100(1-.05) = 95\%$. Examination of Figure 5 leads to an identical conclusion.

Figure 6 portrays a slightly different situation in that Q_3 contains only that portion of $\mu_{X13} - \mu_{X23} = 0$ such that $0 \leq \mu_{X13} \leq 3$ and $0 \leq \mu_{X23} \leq 3$. However, G_3 also contains portions of the same intervals; hence, the conclusion is made that $H_0: \mu_{X13} - \mu_{X23} = 0$ is not rejected. Figure 7 also indicates that $H_0: \mu_{X14} - \mu_{X24} = 0$ might also not be rejected for $\mu_{X14} \in [A_1, B_1]$ and $\mu_{X24} \in [A_2, B_2]$; however, establishing the region G_4 indicates that this conclusion is not tenable and $H_0: \mu_{X14} - \mu_{X24} = 0$ is rejected.

To determine the effect of including the second order term in equations (3.34a), (3.34b) and (3.34c) to obtain a better approximation to Q_4 and G_4 , computations similar to those used for the first order approximation were used to reconstruct the regions. Results indicated that Q_4 was expanded slightly and remained of the same form. The region G_4 was reduced slightly horizontally and expanded vertically so that $\mu_{X14} - \mu_{X24} = 0$ was still contained within G_4 ; hence, conclu-

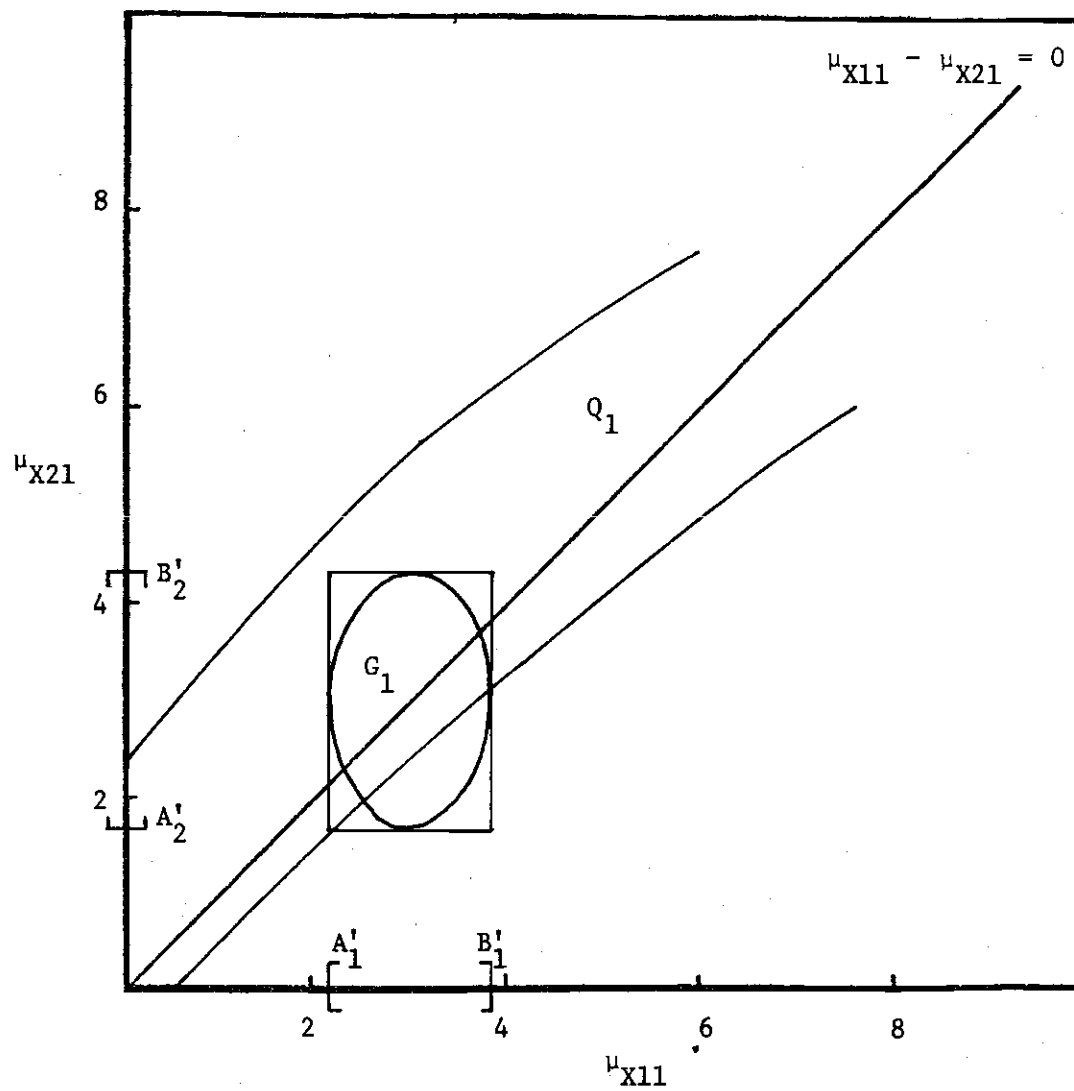


Figure 4. Q and G Regions for $(\mu_{X11} - \mu_{X21})$

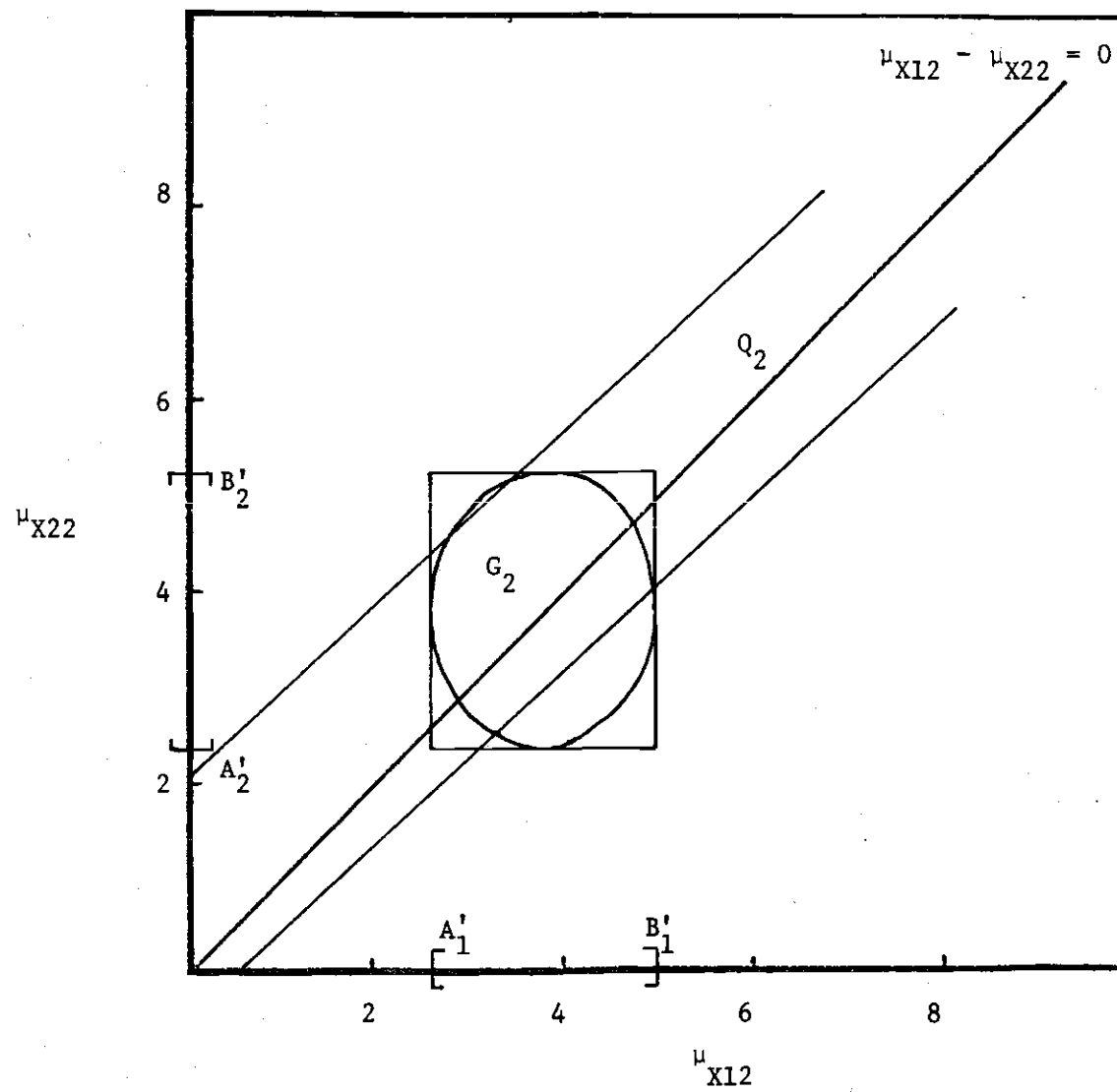


Figure 5. Q and G Regions for $(\mu_{X12} - \mu_{X22})$

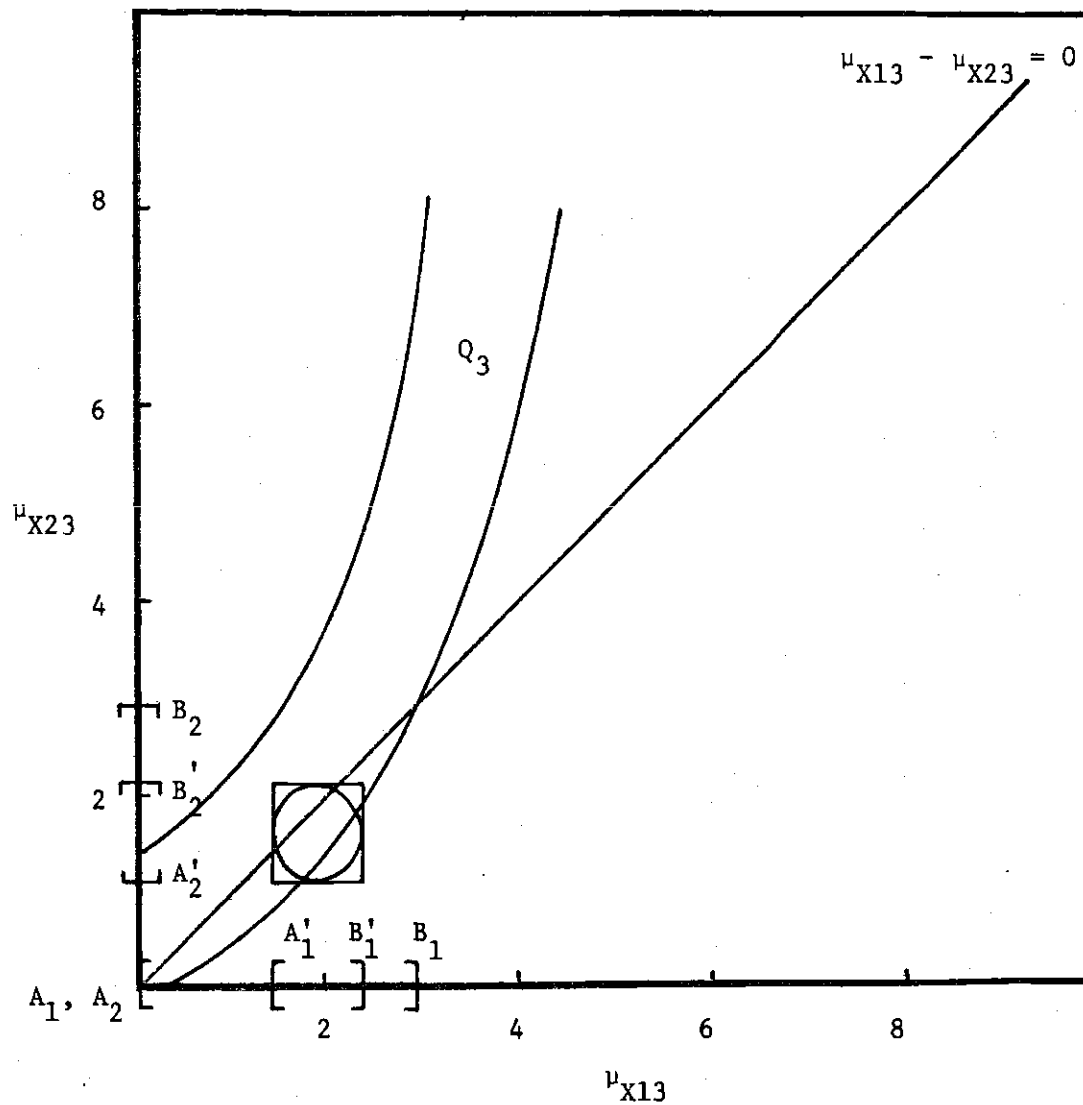


Figure 6. Q and G Regions for $(\mu_{X13} - \mu_{X23})$

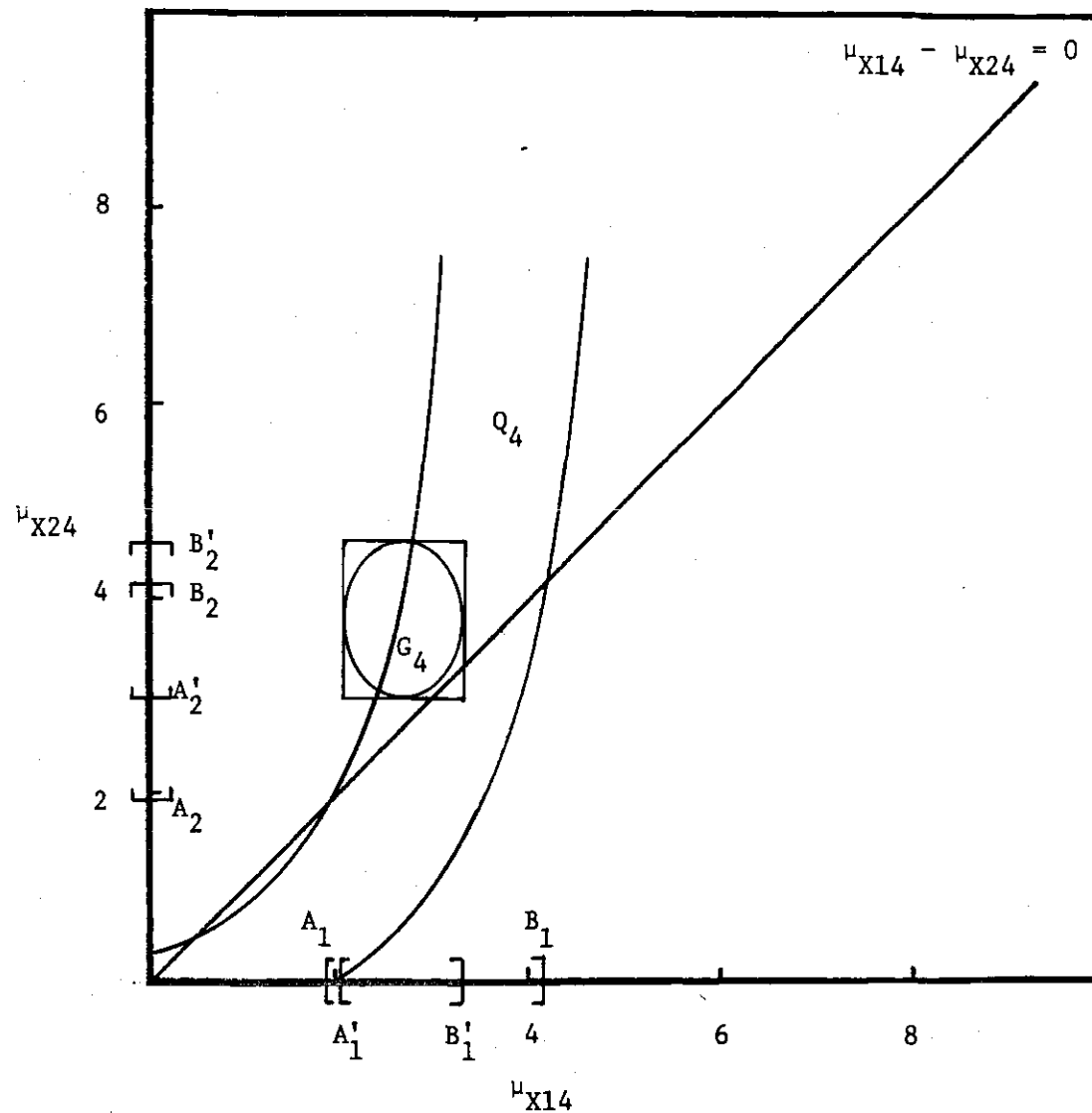


Figure 7. Q and G Regions for $(\mu_{X14} - \mu_{X24})$

sions were the same. Situations may occur where inferences drawn after inclusion of second or higher order terms are different than when only a first order approximation is used. As mentioned previously, then, if transformation parameters, $\hat{\lambda}_{1j}$ and $\hat{\lambda}_{2j}$, are substantially different, a second or higher approximation should probably be used to protect against drawing faulty inferences. Additionally, note should be taken of the fact that inferences based on the graphical procedure are identical to those drawn on the transformed means as depicted in Table 4. Again, this is attributed to the fact that the transformations do not cause much disturbance in the original data.

Further sensitivity of the graphical procedure to other combinations of means and transformation parameters may also be investigated. For purposes of this research, however, sufficient examples have been examined to investigate the problem at hand so that returning to the methodology is appropriate.

Since one of the means of the original sets of MOE appear to be different from each other (MOE_4), and some of the means of MOE describing operational effectiveness of the force using CFV-1 and CFV-2 appear to be different from the standard means, the COEA Project Officer has approved further analysis to determine exact contributions of individual MOE to operational effectiveness. Specifically, note has been taken of the fact that the relative momentum (MOE_4) for the force using CFV-1 and for the force using CFV-2 appear to be different from the standard measure of relative momentum and from each other.

Section II

Extracting principal components from the sample covariance ma-

trices of the original data sets, \underline{X}_1 and \underline{X}_2 , and from $(\underline{X}_1 - \underline{X}_2)$ gives the results in Tables 5-10. Computations for obtaining the principal components and their respective percent contributions, Tables 5, 7, and 9, were performed by a prepared subroutine in (43). The contributions of individual MOE were calculated in the program prepared by the author in Appendix B based on

$$W_i = \sum_{j=1}^k p_j a_{ij}^2,$$

introduced in Chapter III as equation (3.44a).

In examining force operational effectiveness with CFV-1 from Tables 5 and 6, the major concern is with conclusions about MOE_4 , relative momentum. From Table 5, the effect of MOE_4 on each of the principal components is seen to be negative, while its contribution from Table 6 is 21.3847%.

Table 5. Principal Components for \underline{X}_1

Component	Coefficients for $\underline{X}_{1j}(a_{1j})$				Variance Contribution (%)
Y_1	.3838	-.6682	.3110	-.5563	66.9579
Y_2	-.3780	.2581	-.3997	-.7942	21.0570
Y_3	.7376	.6371	.1058	-.1972	11.6998
Y_4	-.4073	.2846	.8558	-.1443	.2853

A similar examination of Tables 7 and 8 for CFV-2, reveals that the effect of relative momentum is also negative in each principal component and has an individual contribution of 23.7283%.

Table 6. MOE Contributions for CFV-1(%)

MOE	Contribution
1	20.4853
2	13.0167
3	45.1134
4	21.3347

Table 7. Principal Components for X_2

Component	Coefficients for X_{2i} (a_{ij})				Variance Contribution (%)
Y_1	.3416	.2260	-.6353	-.6548	75.4756
Y_2	-.4030	-.0761	.5250	-.7458	15.4047
Y_3	.7478	-.5680	.3226	-.1189	9.0207
Y_4	-.4023	-.7877	-.4656	-.0300	.0990

Table 8. MOE Contributions for CFV-2(%)

MOE	Contribution
1	13.2747
2	14.8853
3	48.1117
4	23.7283

For the differences between operational effectiveness with CFV-1 and CFV-2, examination of Tables 9 and 10 shows that differences in relative momentum also have a negative effect and that the contribution of the differences is 24.4449%.

The precise interpretation of these results depends heavily on the definition of other MOE being examined. Additionally, determination of whether contributions are "good" or "bad" depends upon incorporation of the results into the cost portion of the COEA. This also applies to

Table 9. Principal Components for $(X_1 - X_2)$

Components	Coefficients for $(X_{1i} - X_{2i})$				Variance Contribution (%)
Y_1	-.3624	-.4170	-.4828	-.6795	71.4371
Y_2	.3967	.1245	.5803	-.7003	17.8695
Y_3	-.7249	.6482	.1847	-.1424	10.5818
Y_4	.4310	.6249	-.6294	-.1662	.1116

Table 10. MOE Contributions for Differences (%)

MOE	Contribution
1	15.0078
2	15.1391
3	45.4082
4	24.4449

an interpretation of which is the better vehicle. For example, if the scenario-environment combination being examined is battalion in the attack in Environment I, the conclusion might be made that CFV-1 and CFV-2 are not different, with respect to operational effectiveness, since the effect of relative momentum of the force remains approximately the same.

Example Problem II

The structure of this problem is exactly the same as that of Problem I, except that operational mobility will now be measured in terms of five operational performance characteristics, MOE_4 through MOE_8 , common to CFV-1 and CFV-2. Procedurally, the methodology is applied in the same manner with only minor computational adjustments.

In the initial assessment of multivariate normality of X_1 and X_2 ,

the W-statistics and associated confidence levels are:

$$W_1 = .7493 \quad (1-\alpha) < 1\%,$$

and

$$W_2 = .7170 \quad (1-\alpha) < 1\%.$$

Both data sets must therefore be transformed, giving

$$\hat{\lambda}_1 = [.92, .86, .53, .80, 1.31, 1.22, .95, 2.42]; (1-\alpha) = .3194,$$

and

$$\hat{\lambda}_2 = [1.07, 1.13, .95, .89, .65, 1.82, 1.85, .65]; (1-\alpha) = .2071.$$

Re-testing the transformed variables with the W-statistics gives

$$W_1 = .7566 \quad (1-\alpha) < 1\%$$

and

$$W_2 = .7778 \quad (1-\alpha) < 1\%$$

Although the W-statistics have been improved, the confidence levels are still less than 1%, possibly due to the presence of an increased number of outliers in the larger data sets. However, since positive indication of multivariate normality has been achieved based on the likelihood statistic, $\underline{Y}_1 \sim N(\underline{\mu}_{Y1}, \underline{\Sigma}_{Y1})$ and $\underline{Y}_2 \sim N(\underline{\mu}_{Y2}, \underline{\Sigma}_{Y2})$ will be assumed.

Since both $\underline{\Sigma}_{Y1}$ and $\underline{\Sigma}_{Y2}$ are of dimension $p = 8$, Mb^{-1} must be calculated from equations (3.14) and (3.18) to test $H_0: \underline{\Sigma}_{Y1} = \underline{\Sigma}_{Y2}$; hence,

$$M = 30.5366$$

and

$$b^{-1} = .0082$$

giving $Mb^{-1} = .2504$. Comparing this to $F_{.05; f_1, f_2} = 1.312$ reveals that the null hypothesis is not rejected at the 95% level.

$$\text{Testing } H_{01}: \underline{\mu}_{Y1} - \underline{\mu}_0^T = 0, H_{02}: \underline{\mu}_{Y2} - \underline{\mu}_0^T = 0 \text{ and}$$

$H_{03}: \mu_{Y1} - \mu_{Y2} = 0$ gives

$$T_1^2 = 731.3103 > T_{01}^2 = 36.0815: \text{ reject } H_{01},$$

$$T_2^2 = 1421.8828 > T_{02}^2 = 36.0815: \text{ reject } H_{02},$$

and

$$T_3^2 = 239.5794 > T_{03}^2 = 2.7641: \text{ reject } H_{03}.$$

Computationally, equations (3.20) and (3.21) were used to obtain T_1^2 and T_{01}^2 since $H_0: \Sigma_{Y1} = \Sigma_{Y2}$ was not rejected. The conclusions are similar to those in Problem I. Hence, simultaneous confidence intervals will be established about detected differences.

Using equations (3.30) and (3.37) and the approximation procedure presented in Chapter III, intervals on the original means, $(\mu_{X1j} - \mu_{01})$ and $(\mu_{X2j} - \mu_{0j})$, were calculated and are displayed in Table 11. Thus, the original means of MOE_5 and MOE_8 for CFV-1 and MOE_3 , MOE_6 , and MOE_7

Table 11. Confidence Intervals for $(\mu_{Xij} - \mu_{0j})$

CFV-1	CFV-2
$-2.0416 \leq \mu_{X11} - \mu_{01} \leq .3800$	$-3.1015 \leq \mu_{X21} - \mu_{01} \leq .9300$
$-2.1020 \leq \mu_{X12} - \mu_{02} \leq 1.9751$	$-3.1897 \leq \mu_{X22} - \mu_{02} \leq .3934$
$-.7836 \leq \mu_{X13} - \mu_{03} \leq 1.7978$	$-1.8422 \leq \mu_{X23} - \mu_{03} \leq -.4176$
$-2.5757 \leq \mu_{X14} - \mu_{04} \leq .4149$	$-1.1544 \leq \mu_{X24} - \mu_{04} \leq 1.1964$
$-2.8748 \leq \mu_{X15} - \mu_{05} \leq -.9381$	$-2.3009 \leq \mu_{X25} - \mu_{05} \leq 2.1324$
$-2.0978 \leq \mu_{X16} - \mu_{06} \leq .6438$	$-2.1702 \leq \mu_{X26} - \mu_{06} \leq -.1767$
$-1.9764 \leq \mu_{X17} - \mu_{07} \leq .1173$	$-2.6503 \leq \mu_{X27} - \mu_{07} \leq -1.4464$
$-3.1990 \leq \mu_{X18} - \mu_{08} \leq -1.6262$	$-1.0522 \leq \mu_{X28} - \mu_{08} \leq 2.0587$

for CFV-2 are different from the respective elements of the standard mean vector, μ_0 .

From equation (3.29) the intervals about the differences of the means, $(\mu_{Y1j} - \mu_{Y2j})$, were calculated and are displayed in Table 12. In graphing the Q and G regions by the same procedure used for Example Problem 1, results indicated that the differences between MOE_3 and MOE_8

Table 12. Confidence Intervals for $(\mu_{Y1j} - \mu_{Y2j})$

$$-3.0084 \leq \mu_{Y11} - \mu_{Y21} \leq 1.0017$$

$$-3.7641 \leq \mu_{Y12} - \mu_{Y22} \leq .6532$$

$$-1.6633 \leq \mu_{Y13} - \mu_{Y23} \leq -.4371$$

$$-2.2735 \leq \mu_{Y14} - \mu_{Y24} \leq -.0335$$

$$-1.3715 \leq \mu_{Y15} - \mu_{Y25} \leq 1.6514$$

$$-3.5040 \leq \mu_{Y16} - \mu_{Y26} \leq 1.6587$$

$$-2.6168 \leq \mu_{Y17} - \mu_{Y27} \leq .7600$$

$$5.0169 \leq \mu_{Y18} - \mu_{Y28} \leq 18.5883$$

were significant as determined by an examination of Figures 8-15. In comparing these results to the results obtained by examining Table 12, which indicate that MOE_3 , MOE_4 , and MOE_8 are significantly different, it may be seen that the inference drawn about the differences between MOE_4 is different. This exhibits the fact that inferences drawn after non-linear transformation may be affected by the transformation. Since the means of some of the MOE are different from the standard mean vector and two MOE appear to be different from each other, further investigation is warranted.

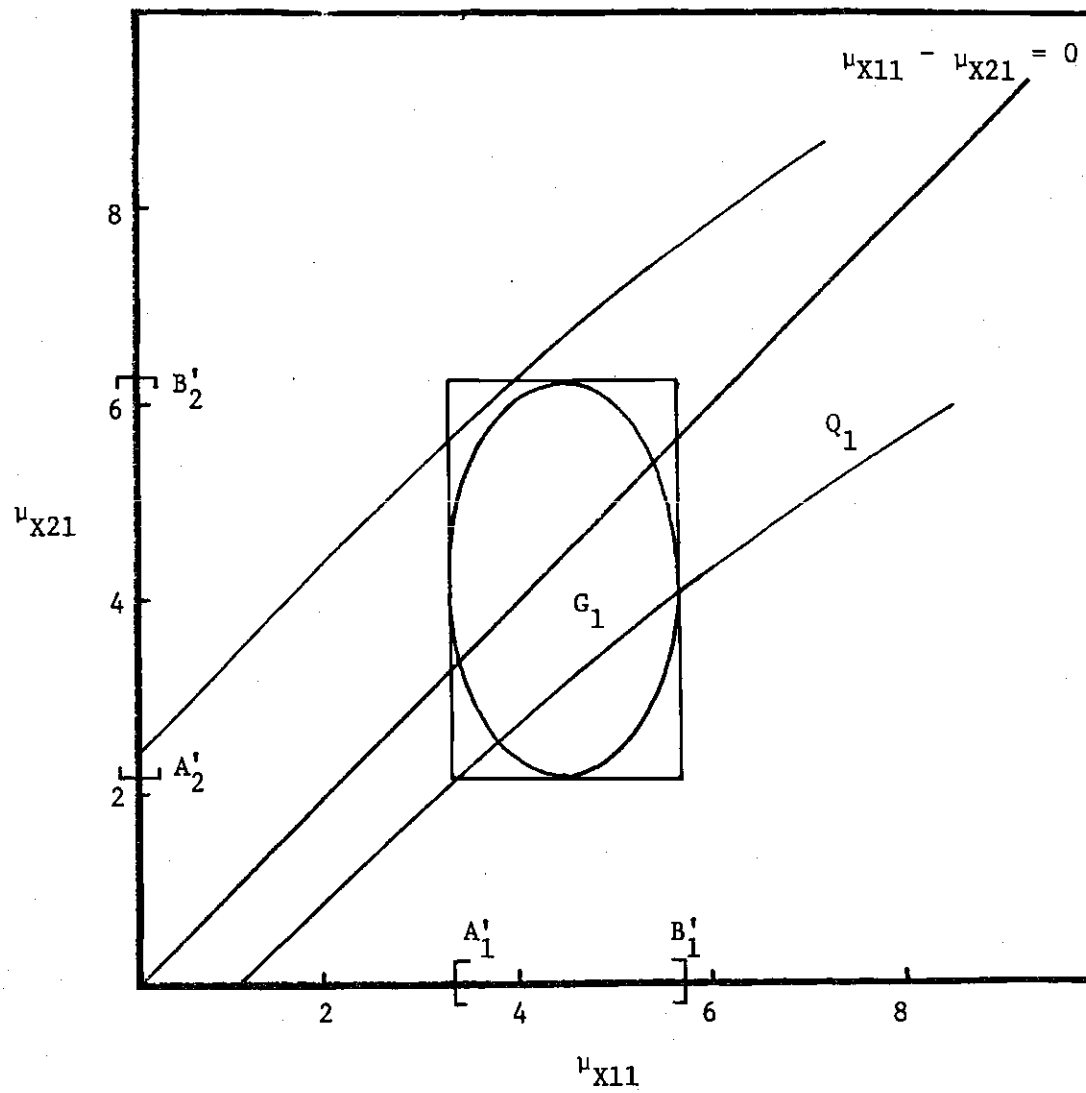


Figure 8. Q and G Regions for $(\mu_{X11} - \mu_{X21})$

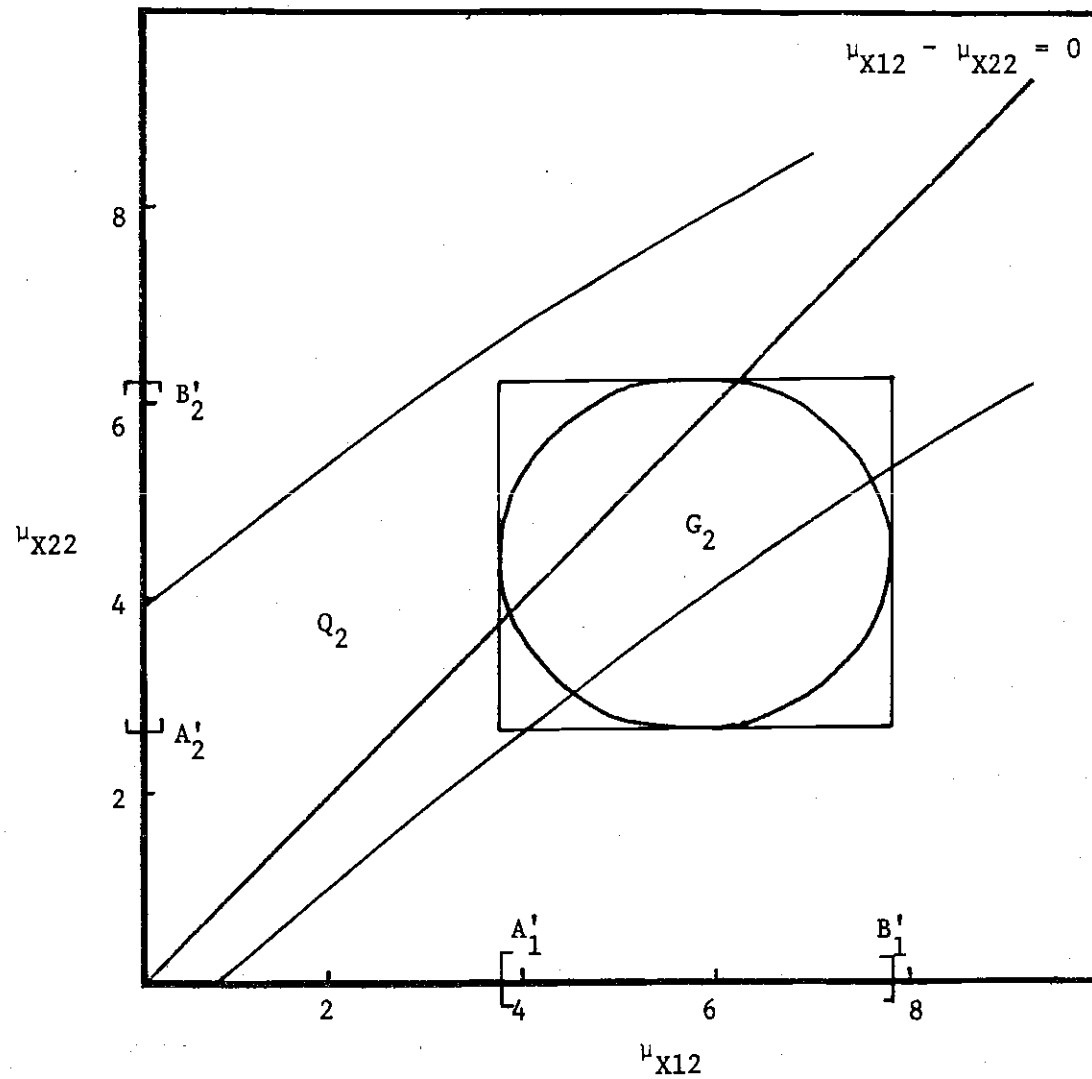


Figure 9. Q and G Regions for $(\mu_{X12} - \mu_{X22})$

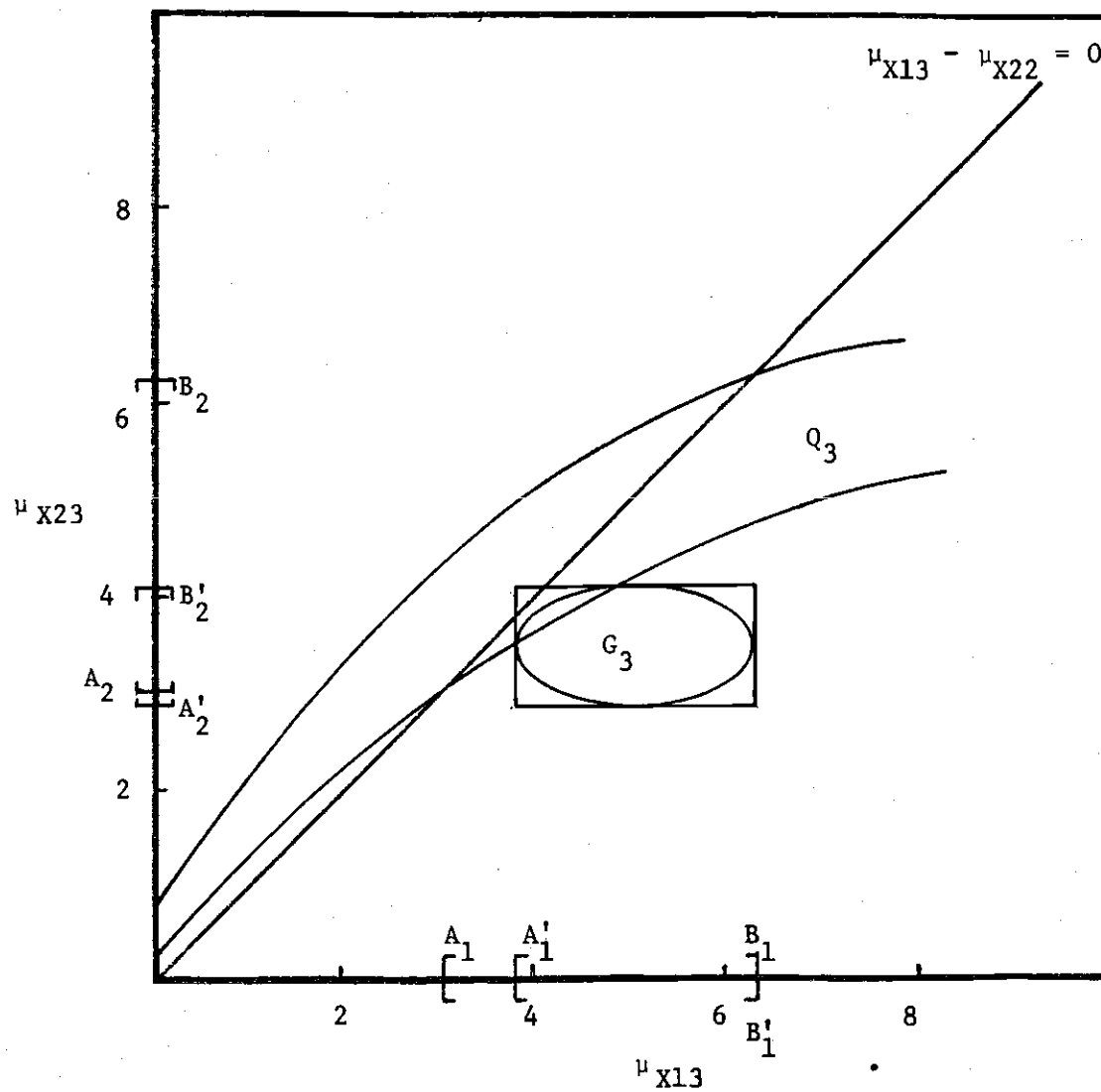


Figure 10. Q and G Regions for $(\mu_{X13} - \mu_{X23})$

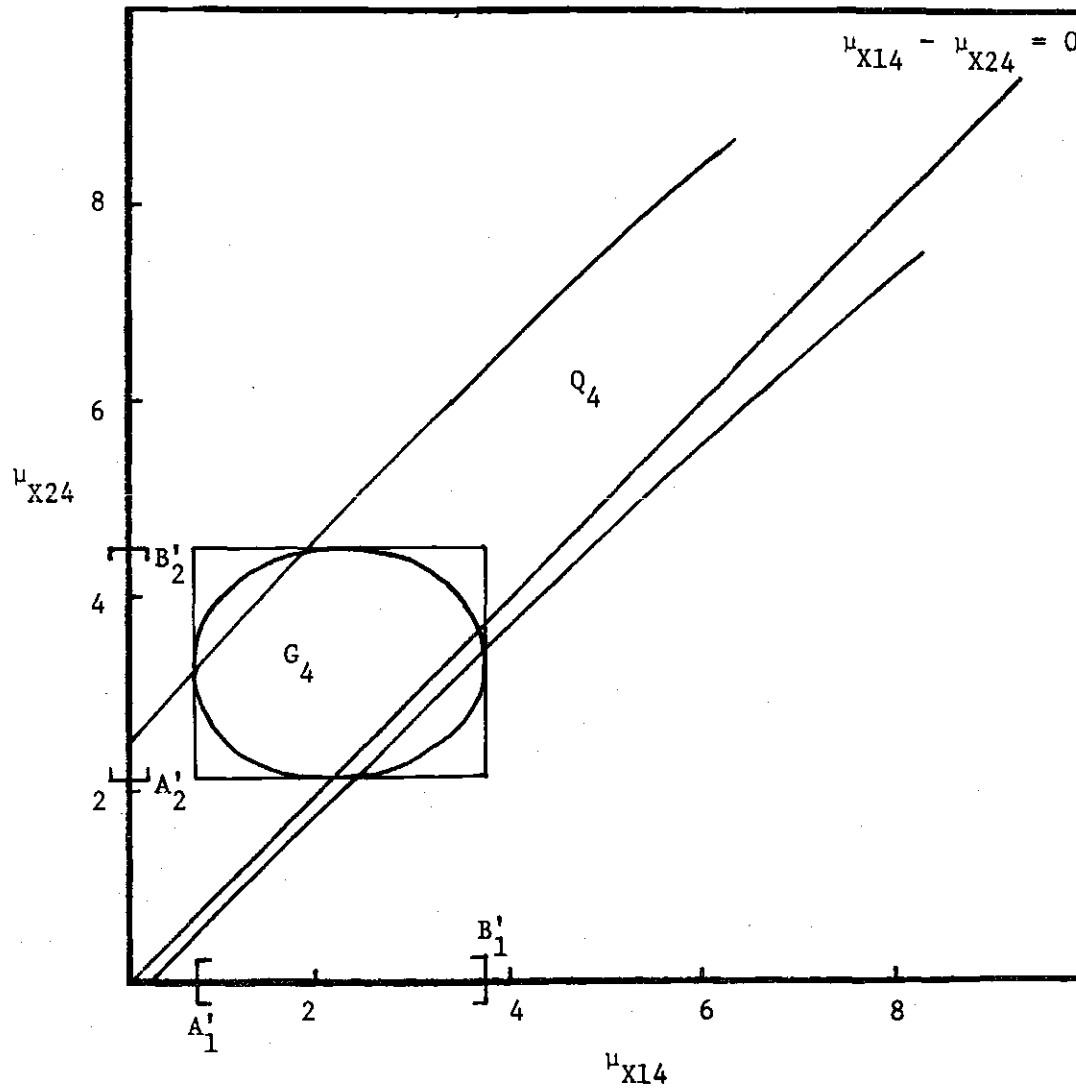


Figure 11. Q and G Regions for $(\mu_{X14} - \mu_{X24})$

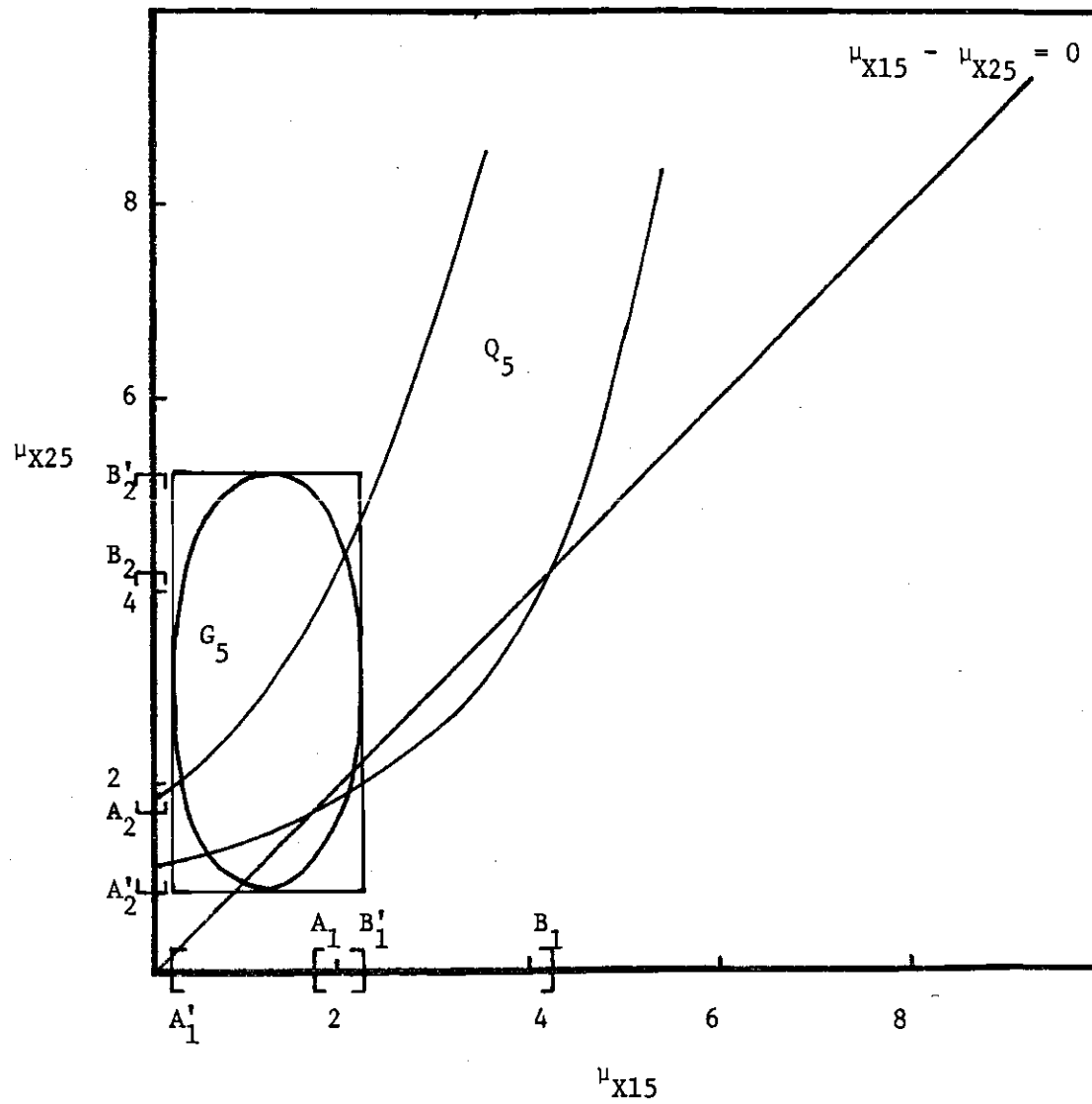


Figure 12. Q and G Regions for $(\mu_{X15} - \mu_{X25})$

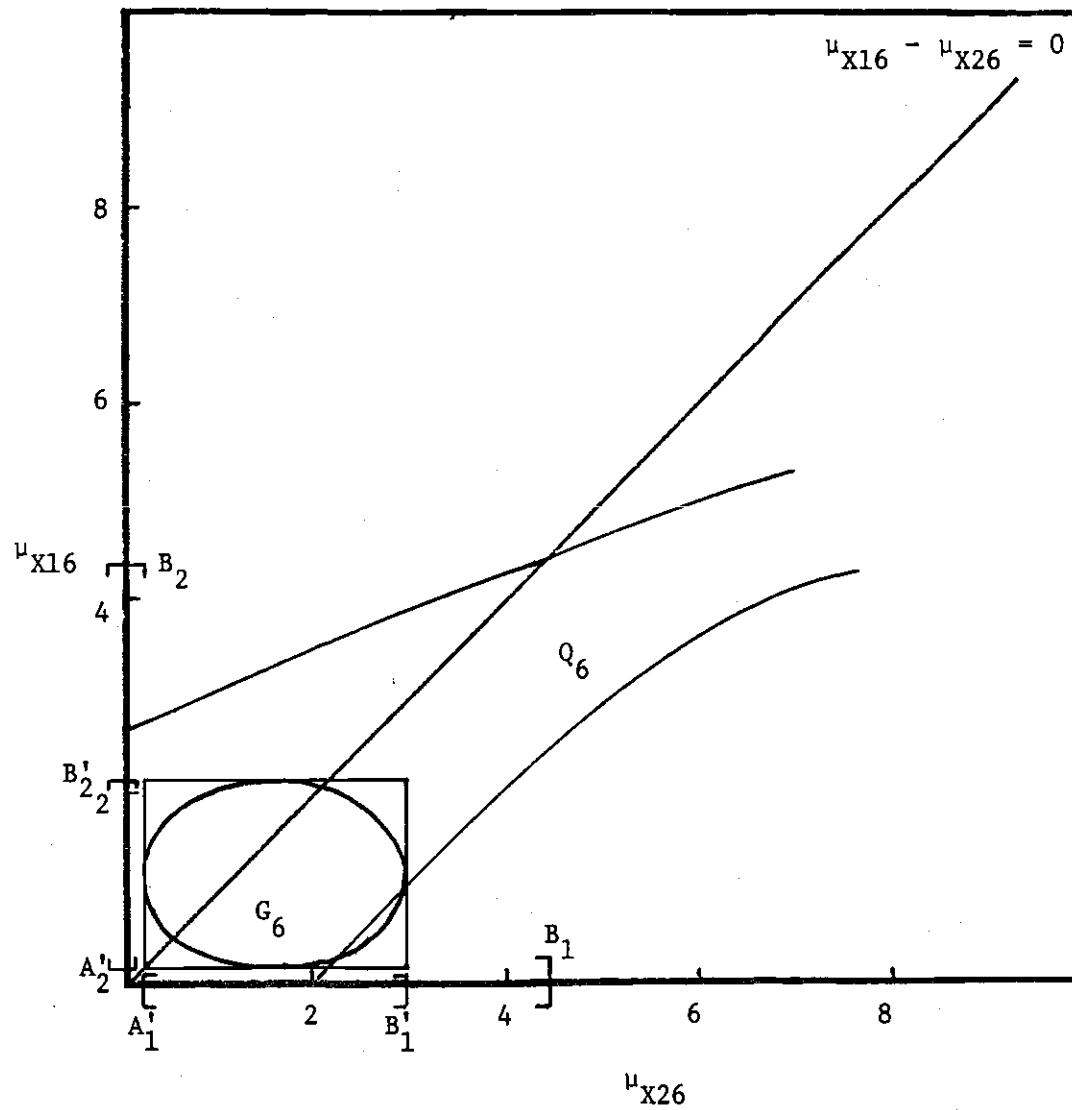


Figure 13. Q and G Regions for $(\mu_{X16} - \mu_{X26})$

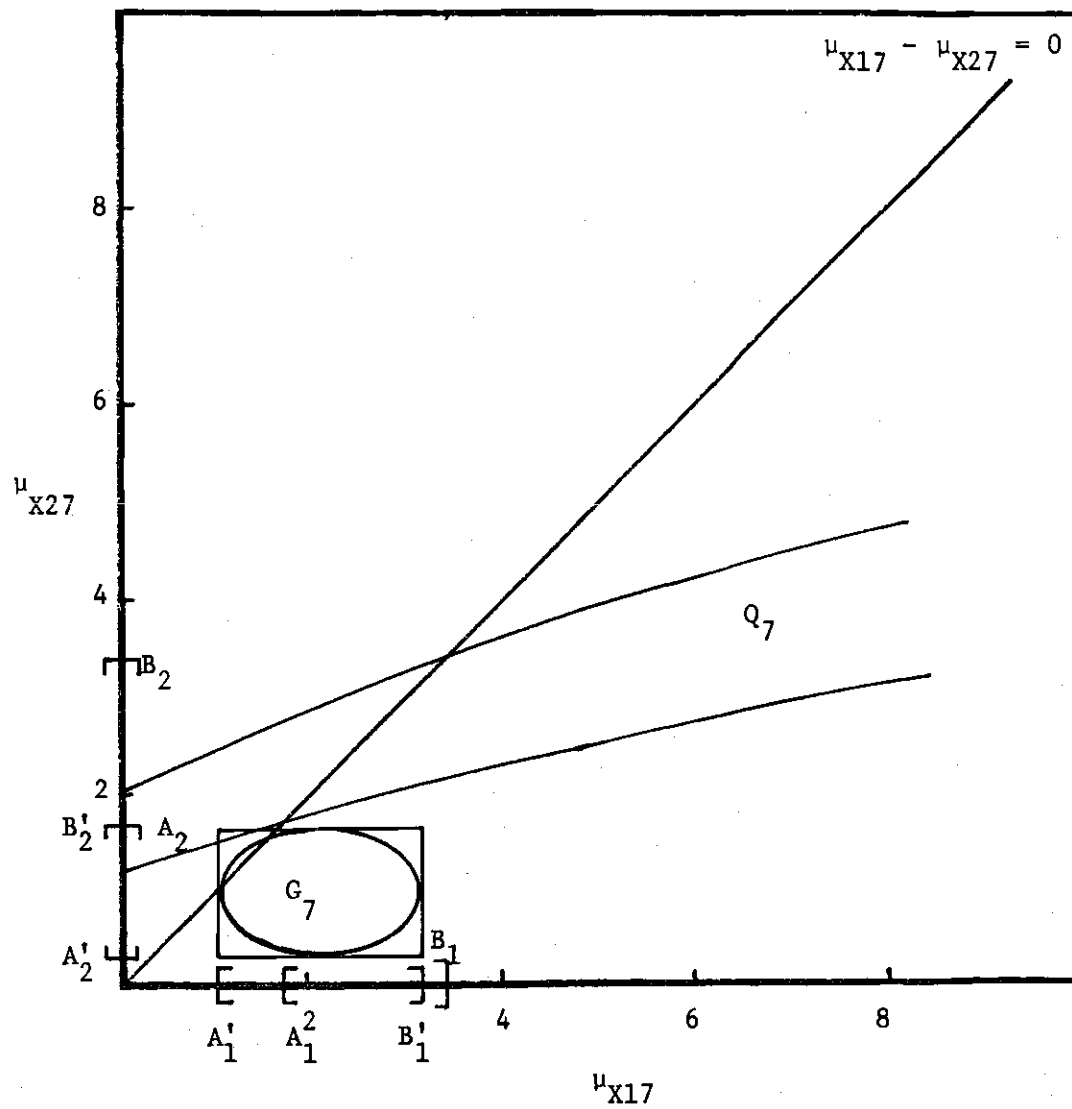


Figure 14. Q and G Regions for $(\mu_{X17} - \mu_{X27})$

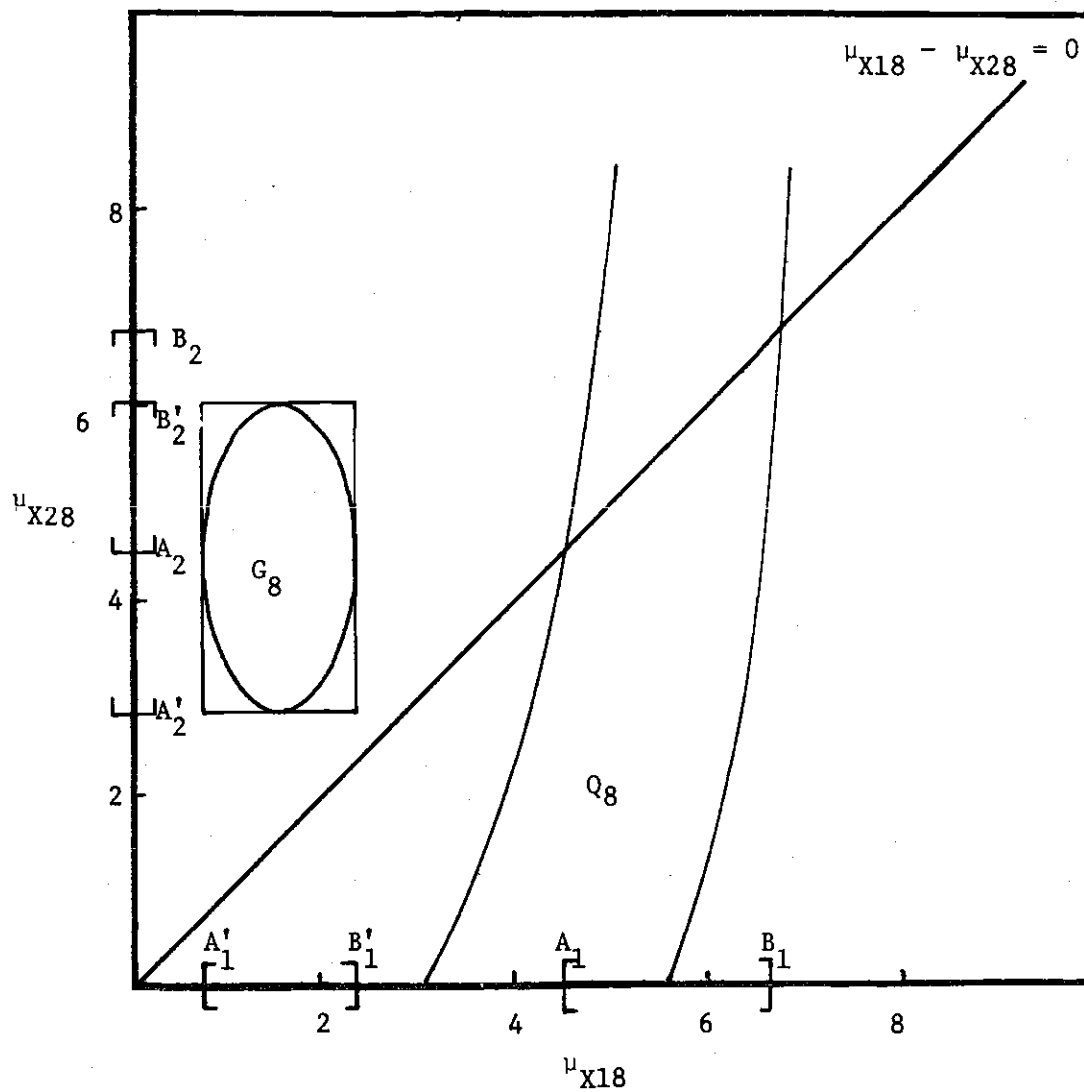


Figure 15. Q and G Regions for $(\mu_{X18} - \mu_{X28})$

The results of the principal component analyses are given in Tables 13-19. Interpretation of these results follows the same procedure as that used for Problem I, with one additional consideration. Since there are now five MOE describing force operational mobility for CFV-1 and CFV-2, their individual contributions must be summed to give the total contribution to operational effectiveness. These results are given in Table 13.

Table 13. Operational Mobility Contributions (%)

	Contribution
CFV-1	51.2433
CFV-2	47.7782
Differences	27.8938

Further Analysis of MOE Contributions

As mentioned previously in the examination of individual MOE contributions in Example Problem I, precise interpretation of percentage contributions is relatively meaningless without knowledge of the nature of MOE being examined. A general analysis may be conducted, however, based on an interpretation of what the percent contributions actually represent.

Recall that the percent contribution of a principal component, Y_i , to the variance within a given data set represents the importance of that linear combination of variables to the variance of the system. By translating the contributions of principal components into contributions of original MOE, the actual result is a determination of the contribution of the variance of the MOE to the variance of the system. In com-

Table 14. Principal Components for X_1

Component	Coefficients for X_{1i} (a_{ij})								Variance Contribution (%)
Y_1	.4058	.7097	-.0849	-.0375	.2282	-.2525	-.1445	.4317	36.9917
Y_2	-.3930	-.0364	.0902	.1867	-.1606	.2655	.2718	.7944	19.4470
Y_3	.7264	-.4463	-.0532	.0474	-.4472	-.0495	.0120	.2559	15.0358
Y_4	-.3743	.0407	-.3685	-.0364	-.5326	-.4295	-.4976	.0732	13.0186
Y_5	-.0702	-.4333	-.3279	-.5797	.4837	-.2202	.0166	.2847	6.6953
Y_6	-.0614	.1154	.0098	-.0673	-.2043	-.5864	.7611	-.1161	6.2591
Y_7	-.0006	.2197	.3528	-.7838	-.3791	.2628	-.0078	-.0079	2.5480
Y_8	-.0659	-.2118	.7835	.0718	.1380	-.4664	-.2792	.1312	.0046

Table 15. Principal Components for X_2

Component	Coefficients for X_{2i} (a_{ij})									Variance Contribution (%)
Y_1	.3754	.3853	-.0062	.1529	.2707	-.0694	-.3352	-.7048		47.1976
Y_2	-.4138	-.3507	.1650	-.0133	-.3142	.4254	.0952	-.6242		16.8396
Y_3	.7154	-.4396	-.1312	-.4607	-.2230	.0730	.0623	-.0806		13.1298
Y_4	-.4051	.1215	-.4274	-.6937	.0582	-.3178	-.1570	-.1678		9.6461
Y_5	-.0252	.0467	-.0166	-.0904	-.1375	.4974	-.8047	.2735		7.4735
Y_6	-.0520	-.3360	-.1383	-.0657	.8473	.3679	.0835	.0251		4.1913
Y_7	.0066	-.0945	-.8604	.4568	-.1809	.0847	.0322	-.0347		1.5182
Y_8	-.0926	-.6297	.1152	.2486	.0724	-.5638	-.4410	-.0476		.0040

Table 16. Principal Components for ($X_1 - X_2$)

Component	Coefficients for ($X_{1i} - X_{2i}$)									Variance Contribution (%)
Y_1	-.3004	-.3716	-.5615	-.0350	-.2090	.6033	-.1891	.1084		70.5931
Y_2	.2906	.0192	.5874	.2102	.0469	.5903	-.4133	.0662		11.0743
Y_3	-.8225	.4881	.2597	-.0129	-.0636	.0814	-.0814	.0167		8.5626
Y_4	.3850	.7386	-.3338	-.2403	-.3099	.1924	-.0447	.0462		4.0911
Y_5	.0144	.1024	-.1251	.9003	-.3686	-.0847	.1377	.0350		3.0826
Y_6	.0122	.0390	.1253	-.0153	.1550	.4577	.8627	.0680		1.4144
Y_7	-.0082	.2565	-.3597	.2931	.8276	.0898	-.1449	-.0697		1.1608
Y_8	.0074	.0085	.0039	-.0090	.0964	-.1386	-.0226	.9853		.0212

Table 17. MOE Contributions for CFV-1 (%)

MOE	Contribution	MOE	Contribution
1	16.8151	5	11.6958
2	7.1215	6	4.3665
3	24.8200	7	12.2033
4	10.9597	8	12.0180

Table 18. MOE Contributions for CFV-2 (%)

MOE	Contribution	MOE	Contribution
1	10.1173	5	2.3111
2	12.0213	6	8.2650
3	30.0832	7	12.1604
4	15.5224	8	9.5193

Table 19. MOE Contributions for Differences (%)

MOE	Contribution	MOE	Contribution
1	11.2972	5	4.0312
2	9.8011	6	1.3971
3	51.0079	7	4.3399
4	18.0435	8	.0821

paring the contributions of MOE to operational effectiveness for the force using CFV-1 (Table 6) with the variances in Σ_{X1} (Appendix B), it may be seen that the order of contributions, high to low, corresponds to the magnitude of variance, low to high. In other words, the MOE that has the highest percent contribution to the variance of the system (MOE₃: 45.1134%) has the lowest variance within the data ($\Sigma_{133} = .4$). This was found to be the case for both data sets in both example problems when the variances were substantially different. When variances were relatively close, then percent contributions were also relatively close, although the order was sometimes reversed.

By saying that a particular MOE contributes a high percent contribution to the variance of a system, then, is equivalent to stating that the variance of that MOE is relatively low when compared to the variances of other MOE. This in turn means that the realization of the random variable representing that particular MOE with low variance is more realistic than a realization of a MOE with high variance. In further analysis, then, a MOE with a high percent contribution, or importance, should be weighted by its importance relative to the other MOE. This is precisely the information provided by determining contributions of individual MOE to the variance of the system and, as mentioned previously, indicates weights of importance of the MOE that may be used in consolidating these results with cost data or other analyses.

CHAPTER VI

SUMMARY AND RECOMMENDATIONS

This research has addressed the analysis of the operational effectiveness of a military force through the use of well-known multivariate statistical techniques applied to hypothesized results of stochastic simulations. Problems examined were restricted to the two sample case with equal sample sizes and population covariance matrices unknown. The overall objective was to develop an improved methodology to determine the contribution of individual MOE to the operational effectiveness of a force, which was accomplished by extending the methods of principal component analysis developed by Hotelling (19) and others.

Several intermediate results obtained during the methodological development were extremely beneficial and deserve special mention. In Section I of the methodology, two tests for multivariate normality were employed in an effort to assess the normality of each set of random variables describing operational effectiveness. The maximum likelihood test developed by Andrews, *et al.*, appears to be more useful since a set of transformation parameters may be obtained that induces multivariate normality in the data. The confidence level associated with a given set of parameters may also be determined. The degree of normality obtained, however, is highly dependent on the efficiency of the particular non-linear optimization procedure selected. The cyclic coordinate method was adopted for reasons of expediency, although a reduced gradient algorithm may have been more efficient for the objective function examined (35).

In addition, the test developed by Malkovich and Afifi appears to perform well only when the underlying distribution is unimodal (27).

Although use of the method of Andrews, *et al.*, eliminates the requirement for ascertaining the exact multivariate distributions of variables under consideration and permits the use of classical T^2 statistics for hypothesis testing, the non-linear transformation eliminated the possibility of drawing inferences directly from transformed data. To alleviate this situation, an approximate graphical procedure was developed in Chapter III, the accuracy of which depends on the magnitude of distribution parameters, the nature of transformation parameters, and the order of approximation used in the Taylor series expansion of the expected value of a function of a random variable. As discussed in Chapter V, including the second term of the expansion did not affect the inferences drawn on the means of the original MOE, but did increase the size of the confidence region slightly.

Section II of the methodology was a straightforward extension of principal component analysis and provides a method for determining the relative importance of MOE being examined. As mentioned previously, precise interpretation of the results of Section II is highly dependent on the specific definitions of MOE being examined. In light of this, only general interpretations of the percent contributions were made herein due to the fact that actual data were not available for the example problems.

With respect to the definitions of operational mobility presented in Chapter II, the concept of relative momentum appears to be more practical and useful. This concept significantly reduces the dimensionality

of the problem and would be relatively simple to measure, when compared to Definition 2. The second definition also depends on a subjective determination of performance characteristics used in describing operational mobility. For example, if several, say, three, vehicles are examined and five performance characteristics are measured for each vehicle, then the dimension of the operational effectiveness vector is $p = 15$ without including the other MOE, such as personnel and equipment losses. Although this has no effect on the methodological procedure, the mere magnitude of the problem may make meaningful analysis and interpretation extremely difficult.

With respect to the above concluding statements, then, the author recommends that the methodology presented in this paper be utilized in the analysis of actual data obtained from stochastic simulation of hypothesized combat to determine contributions of MOE being examined. Furthermore, additional research is certainly warranted in several areas previously discussed. Foremost would be the development of a precise computational algorithm to measure relative momentum and/or operational performance characteristics. Direct implementation of the methodology exactly as in Example Problems 1 and 2 would then enable determination of the contribution of operational mobility to the operational effectiveness of a force. Additionally, a more efficient non-linear optimization algorithm might be investigated for the maximization of $L_{\max}(\lambda)$ in conjunction with ascertaining the behavior of the transformation procedure when applied to data with different underlying distributions. Finally, further investigation of the graphical procedure developed in Chapter III should be conducted to determine its sensitivity to varying transformation parameters or entirely different classes of transformations.

APPENDIX A

This appendix contains tables of best linear coefficients, $\{a_{n-i+1}\}$, Table 20, and percentage points of the test statistic, W , Table 21, used for testing multivariate normality of a vector random variable. Both tables have been reproduced in part from reference 39.

Table 20. Best Linear Coefficients, $\{a_{n-i+1}\}$,
for $n = 11, 12, \dots, 20$

i/n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	.3315	.3325	.3325	.3318	.3306	.3290	.3273	.3253	.3232	.3211
3	.2260	.2347	.2412	.2460	.2495	.2521	.2540	.2553	.2561	.2565
4	.1429	.1586	.1707	.1802	.1878	.1939	.1988	.2027	.2059	.2085
5	.0695	.0922	.1099	.1240	.1353	.1447	.1524	.1587	.1641	.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	-	-	.0000	.0240	.0433	.0593	.0725	.0837	.0932	.1013
8	-	-	-	-	.0000	.0196	.0359	.0496	.0612	.0711
9	-	-	-	-	-	-	.0000	.0163	.0303	.0422
10	-	-	-	-	-	-	-	-	.0000	.0140

Table 21. Percentage Points of W for $n = 3, 4, \dots, 20$

n	Confidence Level								
	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
4	.687	.707	.748	.792	.935	.987	.992	.996	.997
5	.686	.715	.762	.806	.927	.979	.986	.991	.993
6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
7	.730	.760	.803	.838	.928	.972	.979	.985	.988
8	.749	.778	.818	.851	.932	.972	.978	.984	.987
9	.764	.791	.829	.859	.935	.972	.978	.984	.986
10	.781	.806	.842	.869	.938	.972	.978	.983	.986
11	0.792	0.817	0.850	0.876	0.940	0.973	0.979	0.984	0.986
12	.805	.828	.859	.883	.943	.973	.979	.984	.986
13	.814	.837	.866	.889	.945	.974	.979	.984	.986
14	.825	.846	.874	.895	.947	.975	.980	.984	.986
15	.835	.855	.881	.901	.950	.975	.980	.984	.987
16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
17	.851	.869	.892	.910	.954	.977	.981	.985	.987
18	.858	.874	.897	.914	.956	.978	.982	.986	.988
19	.863	.879	.901	.917	.957	.978	.982	.986	.988
20	.868	.884	.905	.920	.959	.979	.983	.986	.988

APPENDIX B

This appendix contains a complete FORTRAN IV listing of the computer program developed by the author to implement the methodology presented in Chapter IV. The program is entirely interactive and is preceded by output for Example Problem 1. Input for the program is in free field format.

MULTIVARIATE STATISTICAL ANALYSIS PROGRAM

ENTER ALPHA

? .05

DO YOU WISH TO EXTRACT PRINCIPAL COMPONENTS FROM COVARIANCE MATRICES
(YES) OR FROM CORRELATION MATRICES (NO)?

? YES

DO YOU WISH TO ENTER DATA MATRICES (YES) OR GENERATE OBSERVATIONS (NO)?
?NO

BEFORE GENERATING DATA, YOU MUST INPUT SOME ADDITIONAL INFORMATION

ENTER NUMBER OF MOE BEING EXAMINED

? 4

ENTER NUMBER OF SYSTEMS BEING EXAMINED

? 2

ENTER NUMBER OF REPLICATIONS YOU WISH TO GENERATE FOR EACH SYSTEM

? 20,20

ENTER COVARIANCE MATRIX FOR SYSTEM 1

? 1, .93, .06, .17

? .93, 1.77, .5, .11

? .06, .5, .4, .31

? .17, .11, .31, .69

ENTER COVARIANCE MATRIX FOR SYSTEM 2

? 1.37, 1.15, .08, .18

? 1.15, 1.96, .59, .11

? .08, .59, .49, .32

? .18, .11, .32, .59

ENTER MEAN VECTOR FOR SYSTEM 1

? 3, 4, 1.9, 2.5

ENTER MEAN VECTOR FOR SYSTEM 2

? 3.5, 4.2, 2.1, 2.3

ENTER VECTOR OF STANDARD MEANS

? 3.2, 3.8, 2.4, 2

ENTER SHAPIRO-WILK COEFFICIENTS FOR SYSTEM 1

? .4734, .3211, .3565, .2085, .1686, .1334, .1013, .0711, .0422, .014

ENTER SHPAIRO-WILK COEFFICIENTS FOR SYSTEM 2

? .4734, .3211, .3565, .2085, .1686, .1334, .1013, .0711, .0422, .014

THE INPUT DATA IS AS FOLLOWS:

ALPHA = .05

COVARIANCE MATRIX FOR SYSTEM 1 IS:

1.0000	.9300	.0600	.1700
.9300	1.7700	.5000	.1100
.0600	.5000	.4000	.3100
.1700	.1100	.3100	.6900

COVARIANCE MATRIX FOR SYSTEM 2 IS:

1.3700	1.1500	.0800	.1800
1.1500	1.9600	.5900	.1100
.0800	.5900	.4900	.3200
.1800	.1100	.3200	.5900

MEAN VECTOR FOR SYSTEM 1 IS:

3.0000	4.0000	1.9000	2.5000
--------	--------	--------	--------

MEAN VECTOR FOR SYSTEM 2 IS:

3.5000	4.2000	2.1000	2.3000
--------	--------	--------	--------

VECTOR OF STANDARD MEANS IS:

3.2000	3.8000	2.4000	2.0000
--------	--------	--------	--------

NUMBER OF MOE BEING EXAMINED IS 4

NUMBER OF ALTERNATIVE SYSTEMS IS 2

NUMBER OF REPLICATIONS FOR SYSTEM 1 IS 20

NUMBER OF REPLICATIONS FOR SYSTEM 2 IS 20

SHAPIRO-WILK COEFFICIENTS FOR SYSTEM 1 ARE:

.4734	.3211	.2565	.2085
.1686	.1334	.1013	.0711
.0422	.0140		

SHAPIRO-WILK COEFFICIENTS FOR SYSTEM 2 ARE:

.4734	.3211	.2565	.2085
.1686	.1334	.1013	.0711
.0422	.0410		

PRINCIPAL COMPONENTS WILL BE EXTRACTED FROM THE SAMPLE COVARIANCE MATRICES

IS YOUR DATA CORRECT?

? YES

YOUR DATA MATRICES HAVE BEEN GENERATED AND WILL NOW BE TESTED FOR MULTIVARIATE NORMALITY

THE SHAPIRO-WILK STATISTIC FOR SYSTEM 1 IS: .9317

THE SHAPIRO-WILK STATISTIC FOR SYSTEM 2 IS: .8378

DOES THE DATA FOR SYSTEM 1 PASS YOUR CRITERIA FOR MULTIVARIATE NORMALITY?
? NO

DOES THE DATA FOR SYSTEM 2 PASS YOUR CRITERIA FOR MULTIVARIATE NORMALITY?
? NO

THE DATA FOR BOTH SYSTEMS WILL NOW BE TRANSFORMED AND TESTED AGAIN FOR MULTIVARIATE NORMALITY

TRANSFORMATION PARAMETERS FOR SYSTEM 1

LAMBDA (1) =	.9200
LAMBDA (2) =	.9500
LAMBDA (3) =	1.1000
LAMBDA (4) =	1.5800

LIKLIHOOD RATIO STATISTIC = 3.5359 CONFIDENCE LEVEL = .5276

THE SHAPIRO-WILK STATISTIC FOR SYSTEM 1 IS: .9456

TRANSFORMATION PARAMETERS FOR SYSTEM 2

LAMBDA (1) =	.9800
LAMBDA (2) =	.9800
LAMBDA (3) =	.5900
LAMBDA (4) =	.5900

LIKLIHOOD RATIO STATISTIC = 5.0926 CONFIDENCE LEVEL = .7221

THE SHAPIRO-WILK STATISTIC FOR SYSTEM 2 IS: .8991

DOES THE DATA FOR SYSTEM 1 PASS YOUR CRITERIA FOR MULTIVARIATE NORMALITY?
? YES

SINCE THE TRANSFORMED DATA FOR SYSTEM 1 HAS PASSED THE TEST FOR MULTIVARIATE NORMALITY, THIS TRANSFORMED DATA WILL BE USED IN THE FOLLOWING HYPOTHESIS TESTS

DOES THE DATA FOR SYSTEM 2 PASS YOUR CRITERIS FOR MULTIVARIATE NORMALITY?
? YES

SINCE THE TRANSFORMED DATA FOR SYSTEM 2 HAS PASSED THE TEST FOR MULTIVARIATE NORMALITY, THIS TRANSFORMED DATA WILL BE USED IN THE FOLLOWING HYPOTHESIS TESTS

WE WILL NOW TEST THE SYSTEMS FOR EQUALITY OF COVARIANCE MATRICES

BARTLETT'S STATISTIC = 50.8638 CHI-SQ. STATISTIC = 18.3070

THE NULL HYPOTHESIS IS REJECTED

WE WILL NOW TEST THE MEANS OF THE SYSTEMS

TEST FOR EQUALITY OF SYSTEM 1 MEAN VECTOR AND STANDARD MEAN VECTOR

T-SQ. STATISTIC = 1294.1079 HOTELLING TEST STATISTIC = 14.2829

THE NULL HYPOTHESIS IS REJECTED

TEST FOR EQUALITY OF SYSTEM 2 MEAN VECTOR AND STANDARD MEAN VECTOR

T-SQ. STATISTIC = 669.0139 HOTELLING TEST STATISTIC = 14.2829

THE NULL HYPOTHESIS IS REJECTED

TEST FOR EQUALITY OF SYSTEM 1 AND SYSTEM 2 MEAN VECTORS

T-SQ. STATISTIC = 255.4379 HOTELLING TEST STATISTIC = 14.2829

THE NULL HYPOTHESIS IS REJECTED

SIMULTANEOUS CONFIDENCE INTERVALS WILL NOW BE PLACED ABOUT THE DIFFERENCES OF THE MEANS

SYSTEM 1 MEAN VECTOR AND STANDARD MEAN VECTOR
 LOWER LIMIT UPPER LIMIT

-1.0234	.5563
-1.0693	1.0111
-1.0937	-.1444
.0118	2.1259

SYSTEM 2 MEAN VECTOR AND STANDARD MEAN VECTOR
 LOWER LIMIT UPPER LIMIT

-1.5328	1.0657
-1.4365	1.3783
-1.0009	-.2372
.6030	1.5347

SYSTEM 1 AND SYSTEM 2 MEAN VECTORS
 LOWER LIMIT UPPER LIMIT

-2.6116	.5279
-2.5052	.7205
-.6057	.6010
.2200	2.5489

PRINCIPAL COMPONENTS WILL NOW BE COMPUTED

EIGENVALUES FOR SYSTEM 1 DATA MATRIX

.0104 .4257 .7662 2.4363

EIGENVECTORS FOR SYSTEM 1 DATA MATRIX

.3838	-.6682	.3110	-.5563
-.3780	.2581	-.3997	-.7942
.7376	.6371	.1058	-.1972
-.4073	.2846	.8558	-.1443

PERCENT OF SYSTEM 1 VARIANCE ATTRIBUTED TO ITS PRINCIPAL COMPONENTS

.2853 11.6998 21.0570 66.9579

PERCENT CONTRIBUTION OF INDIVIDUAL MOE TO OPERATIONAL EFFECTIVENESS WITH
 SYSTEM 1

MOE	CONTRIBUTUION
1	20.4853
2	13.0167
3	45.1134
4	21.3847

EIGENVALUES FOR SYSTEM 2 DATA MATRIX

.0063 .5744 .9809 4.8061

EIGENVECTORS FOR SYSTEM 2 DATA MATRIX

.3416	.2260	-.6353	-.6548
-.4030	-.0761	.5250	-.7458
.7478	-.5680	.3226	-.1189
-.4023	-.7877	-.4656	-.0300

PERCENT OF SYSTEM 2 VARIANCE ATTRIBUTED TO ITS PRINCIPAL COMPONENTS

.0990 9.0207 15.4047 75.4756

PERCENT CONTRIBUTION OF INDIVIDUAL MOE TO OPERATIONAL EFFECTIVENESS WITH SYSTEM 2

MOE	CONTRIBUTION
1	13.2747
2	14.8853
3	48.1117
4	23.7283

EIGENVALUES FOR DIFFERENCE MATRIX

.0106 1.0091 1.7040 6.8121

EIGENVECTORS FOR DIFFERENCE MATRIX

-.3624	-.4170	-.4828	-.6795
.3967	.1245	.5803	-.7003
-.7249	.6482	.1847	-.1424
.4310	.6249	-.6294	-.1662

PERCENT VARIANCE OF DIFFERENCE MATRIX ATTRIBUTED TO ITS PRINCIPAL COMPONENTS

.1116 10.5818 17.8695 71.4371

PERCENT CONTRIBUTION OF DIFFERENCES IN MOE TO DIFFERENCES IN OPERATIONAL EFFECTIVENESS

MOE	CONTRIBUTION
1	15.0078
2	15.1391
3	45.4082
4	24.4449

```

PROGRAM MAIN(INPUT,OUTPUT,TAPE3,TAPE5=INPUT,TAPE6=
*  OUTPUT)
COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
COMMON/FOUR/US(2,20),SW(2),A(2,20)
COMMON/FIVE/YBAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SIX/AINVER(4,4),TEMPD(2,4,20),RMAX(2,20)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
COMMON/NINE/EVAL(4),EVEC(4,4),VAR(4),XCONT(4)
COMMON/TEN/CLIML(4),CLIMU(4)
COMMON/ELEVEN/X1(4),XSTAR(4)
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
COMMON/BLOCK1/JPP,MPP,LPP,COUNT
COMMON/BLOCK2/YDIFF(2,4),XT(2,4),SZ1(4,4),SY1(4,4)
DIMENSION LANS(2),JANS(2)
REAL MC
INTEGER COUNT
COUNT=0
KOUNT=0
IMAX=0
IA=1
IB=1
DATA KANS/6HYES /

```

DATA INPUT SECTION

```

C
C
C
WRITE(6,102)
185 WRITE(6,103)
   READ(3,*) ALPHA
   IF(EOF(3)) 1000,190
190 WRITE(6,104)
   READ(3,101) IANS1
   IF(EOF(3)) 1000,195
195 IF(IANS1.EQ.KANS) GO TO 200
   IA=2
200 WRITE(6,105)
   READ(3,101) IANS2
   IF(EOF(3)) 1000,202
202 IF(IANS2.EQ.KANS) GO TO 205
   IB = 2
   GO TO 215
205 WRITE(6,106)
   WRITE(6,107)
   READ(3,*) IP
   IF(EOF(3)) 1000,207
207 WRITE(6,108)
   READ(3,*) NA
   IF(EOF(3)) 1000,208

```

```

208 WRITE(6,109)
   READ(3,*) (N(I),I=1,NA)
   IF(EOF(3)) 1000,206
206 DO 210 I=1,NA
   M=N(I)
   WRITE(6,110) I
210 READ(3,*) ((X(I,J,K),K=1,M),J=1,IP)
   IF(EOF(3)) 1000,203
203 WRITE(6,145)
   READ(3,*) (XMU(I),I=1,IP)
   IF(EOF(3)) 1000,209
209 DO 214 I=1,NA
   MS=N(I)
   IF(MS/FLOAT(MS).EQ.1.) GO TO 211
   KI=(MS+1)/2
   GO TO 212
211 KI=MS/2
212 WRITE(6,129) I
214 READ(3,*) (A(I,J),J=1,KI)
   IF(EOF(3)) 1000,226
215 WRITE(6,111)
   WRITE(6,107)
   READ(3,*) IP
   IF(EOF(3)) 1000,216
216 WRITE(6,108)
   READ(3,*) NA
   IF(EOF(3)) 1000,217
217 WRITE(6,128)
   READ(3,*) (N(I),I=1,NA)
   IF(EOF(3)) 1000,218
218 DO 220 I=1,NA
   WRITE(6,112) I
220 READ(3,*) ((SIGMA(I,J,K),K=1,IP),J=1,IP)
   IF(EOF(3)) 1000,221
221 DO 225 I=1,NA
   WRITE(6,113) I
225 READ(3,*) (SMU(I,J),J=1,IP)
   IF(EOF(3)) 1000,228
228 WRITE(6,145)
   READ(3,*) (XMU(I),I=1,IP)
   IF(EOF(3)) 1000,219
219 DO 224 I=1,NA
   MS=N(I)
   IF(MS/FLOAT(MS).EQ.1.) GO TO 222
   KI=(MS+1)/2
   GO TO 223
222 KI=MS/2
223 WRITE(6,129) I
224 READ(3,*) (A(I,J),J=1,KI)
   IF(EOF(3)) 1000,226

```

```

226 WRITE(6,126)
    WRITE(6,114)
    WRITE(6,115) ALPHA
    IF(IANS2.EQ.KANS) GO TO 230
    GO TO 245
230 DO 235 I=1,NA
    M=N(I)
    WRITE(6,116) I
235 WRITE(6,*) ((X(I,J,K),K=1,M),J=1,IP)
    WRITE(6,117) IP
    WRITE(6,113) NA
    DO 240 I=1,NA
240 WRITE(6,119) I,N(I)
    DO 241 I=1,NA
    WRITE(6,130) I
241 WRITE(6,125) (A(I,J),J=1,KI)
    GO TO 260
245 DO 250 I=1,NA
    WRITE(6,120) I
250 WRITE(6,125) ((SIGMA(I,J,K),K=1,IP),J=1,IP)
    DO 255 I=1,NA
    WRITE(6,121) I
255 WRITE(6,125) (SMU(I,J),J=1,IP)
    WRITE(6,146) (XMU(I),I=1,IP)
    WRITE(6,117) IP
    WRITE(6,118) NA
    DO 256 I=1,NA
256 WRITE(6,119) I,N(I)
    DO 257 I=1,NA
    WRITE(6,130) I
257 WRITE(6,125) (A(I,J),J=1,KI)
260 IF(IA.EQ.2) GO TO 265
    WRITE(6,122)
    GO TO 270
265 WRITE(6,123)
270 WRITE(6,124)
    READ(5,101) IANS3
    IF(EOF(5)) 1000,271
271 IF(IANS3.EQ.KANS) GO TO 277
    GO TO 185
277 IF(IANS2.NE.KANS) GO TO 276
    GO TO 304

```

C
C
C

GENERATE OBSERVATIONS IF DATA NOT AVAILABLE

```

276 CALL GEN
    WRITE(6,131)
    GO TO 305

```

C
C

TEST THE DATA FOR MULTIVARIATE NORMALITY

```

304 WRITE(6,136)
305 IPP=1
    CALL MTEST
    IPP=2
    CALL MTEST
    IPP=0
    CALL OUTPUT(4)
    DO 300 I=1,NA
    WRITE(6,132) I
    READ(5,101) JANS(I)
    IF(EOF(5)) 1000,300
300 CONTINUE

```

```

C
C      IF THE DATA IS NOT MVN, APPLY THE TRANSFORMATION
C      AND TEST AGAIN FOR MVN.
C

```

```

    IF(JANS(1).EQ.KANS.AND.JANS(2).EQ.KANS) GO TO 395
    IF(JANS(1).EQ.KANS.AND.JANS(2).NE.KANS) GO TO 325
    IF(JANS(2).EQ.KANS.AND.JANS(1).NE.KANS) GO TO 331
    WRITE(6,133)
    KOUNT=1
    GO TO 331
325 WRITE(6,134) 2
326 IPP=2
    CALL TRANS(IP)
    CALL OUTPUT(5)
    CALL MTEST
    CALL OUTPUT(4)
    IF(KOUNT.EQ.1) IPP=1
    GO TO 352
330 WRITE(6,134) 1
331 IPP=1
    CALL TRANS(IP)
    CALL OUTPUT(5)
    CALL MTEST
    CALL OUTPUT(4)
    IF(KOUNT.EQ.1) GO TO 326
352 IF(IPP.EQ.1) GO TO 372
    IF(IPP.EQ.2) GO TO 358
358 WRITE(6,132) 2
    READ(5,101) IANS6
    IF(EOF(5)) 1000,363
363 IF(IANS6.EQ.KANS) GO TO 364
    GO TO 371
364 WRITE(6,143) 2
    WRITE(6,126)
    GO TO 360
371 WRITE(6,135) 2
    DO 366 J=1,IP
    DO 367 K=1,M

```

```

367 X(2,J,K)=ALOG(PVAR(2,J,K))
366 CONTINUE
    GO TO 360
372 WRITE(6,132) 1
    READ(5,101) IANS7
    IF(EOF(5)) 1000,383
383 IF(IANS7.EQ.KANS) GO TO 384
    GO TO 391
384 WRITE(6,143) 1
    IF(KOUNT.EQ.1) GO TO 358
    GO TO 360
391 WRITE(6,135) 1
    DO 386 J=1,IP
    DO 387 K=1,M
387 X(1,J,K)=ALOG(PVAR(1,J,K))
386 CONTINUE

```

```

C
C      TEST EQUALITY OF COVARIANCE MATRICES
C

```

```

395 IK=1
360 WRITE(6,137)
    CALL STEST
    WRITE(6,126)

```

```

C
C      TEST MEAN VECTORS AND CALCULATE CONFIDENCE
C      INTERVALS IF DIFFERENCES IN MEANS ARE DETECTED
C

```

```

    WRITE(6,139)
    CALL MUTEST
    IF(KPP.EQ.1) GO TO 390
    WRITE(6,126)
    WRITE(6,141)
    CALL CONINT

```

```

C
C      EXTRACT PRINCIPAL COMPONENTS AND CALCULATE MOE
C      CONTRIBUTIONS
C

```

```

    WRITE(6,126)
390 WRITE(6,142)
    CALL PRICOM
101 FORMAT(A6)
102 FORMAT(20X,*MULTIVARIATE STATISTICAL ANALYSIS
* PROGRAM*)
103 FORMAT(/,*ENTER ALPHA*)
104 FORMAT(/,*DO YOU WISH TO EXTRACT PRINCIPAL COMPONENTS
* FROM COVARIA
*NCE MATRICES(YES)*,/,*OR FROM CORRELATION MATRICES
* ?*)
105 FORMAT(/,*DO YOU WISH TO ENTER DATA MATRICES(YES) OR
* GENERATE *

```



```

**OBSERVATIONS(NO)?*)
106 FORMAT(/,*BEFORE ENTERING DATA, YOU MUST INPUT SOME
*   ADDITIONAL *
**INFORMATION*)
107 FORMAT(/,*ENTER NUMBER OF MOE BEING EXAMINED*)
108 FORMAT(/,*ENTER NUMBER OF SYSTEMS BEING EXAMINED*)
109 FORMAT(/,*ENTER NUMBER OF REPLICATIONS FOR EACH
*   ALTERNATIVE SYSTEM
*, BASE SYSTEM FIRST*)
110 FORMAT(/,*NOW ENTER DATA MATRIX FOR SYSTEM *,I2,/)
111 FORMAT(/,*BEFORE GENERATING DATA, YOU MUST INPUT SOME
*   ADDITIONAL *
**INFORMATION*)
112 FORMAT(/,*ENTER COVARIANCE MATRIX FOR SYSTEM *,I2,/)
113 FORMAT(/,*ENTER MEAN VECTOR FOR SYSTEM *,I2,/)
114 FORMAT(/,21X,*THE INPUT DATA IS AS FOLLOWS:*)
115 FORMAT(/,*ALPHA = *,F4.2)
116 FORMAT(/,*SYSTEM *,I2,* DATA MATRIX IS:*,/)
117 FORMAT(/,*NUMBER OF MOE BEING EXAMINED IS *,I2)
118 FORMAT(/,*NUMBER OF ALTERNATIVE SYSTEMS IS *,I2)
119 FORMAT(/,*NUMBER OF REPLICATIONS FOR SYSTEM *,I2,* IS
*   *,I2)
120 FORMAT(/,*COVARIANCE MATRIX FOR SYSTEM *,I2,* IS:*,/)
121 FORMAT(/,*MEAN VECTOR FOR SYSTEM *,I2,* IS:*)
122 FORMAT(/,*PRINCIPAL COMPONENTS WILL BE EXTRACTED FROM
*   THE SAMPLE
*COVARIANCE MATRICES*)
123 FORMAT(/,*PRINCIPAL COMPONENTS WILL BE EXTRACTED FROM
*   THE SAMPLE
*CORRELATION MATRICES*)
124 FORMAT(/,*IS YOUR DATA CORRECT?*)
125 FORMAT(/,4(1X,F8.4))
126 FORMAT(/,10X,58(1H*))
128 FORMAT(/,*ENTER NUMBER OF REPLICATIONS YOU WISH TO
*   GENERATE *
**FOR EACH SYSTEM*)
129 FORMAT(/,*ENTER SHAPIRO-WILK COEFFICIENTS FOR SYSTEM
*   *,I2,/)
130 FORMAT(/,*SHAPIRO-WILK COEFFICIENTS FOR SYSTEM *,I2,
** ARE:*,/)
131 FORMAT(/,*YOUR DATA MATRICES HAVE BEEN GENERATED AND
*   WILL NOW *
**BE TESTED FOR MULTIVARIATE*,/,*NORMALITY*)
132 FORMAT(/,*DOES THE DATA FOR SYSTEM *,I2,* PASS YOUR
*   CRITERIA *
**FOR MULTIVARIATE NORMALITY?*)
133 FORMAT(/,*THE DATA FOR BOTH SYSTEMS WILL NOW BE
*   TRANSFORMED *
**AND TESTED AGAIN FOR*,/,*MULTIVARIATE NORMALITY*)
134 FORMAT(/,*THE DATA FOR SYSTEM *,I2,* WILL BE
*   TRANSFORMED AND *

```

```

**TESTED AGAIN FOR*,/,*MULTIVARIATE NORMALITY*)
135 FORMAT(/,*SINCE THE TRANSFORMED DATA FOR SYSTEM*,I2,*
* DOES *
**NOT FOLLOW A MULTIVARIATE*,/,*NORMAL DISTRIBUTION,
* EACH MOE *
**WILL BE TRANSFORMED SEPARATELY TO INSURE*,/,
* *MARGINAL NORMAL*
**ITY. THIS DATA WILL THEN BE USED IN THE REMAINDER
* OF*,/,*THE*
** PROGRAM*)
136 FORMAT(/,*YOUR DATA MATRICES WILL NOW BE TESTED FOR
* MULTI*
**VARIATE NORMALITY*)
137 FORMAT(/,*WE WILL NOW TEST THE SYSTEMS FOR EQUALITY
* OF *
**COVARIANCE MATRICES*)
139 FORMAT(/,*WE WILL NOW TEST THE MEANS OF THE SYSTEMS*,/
* )
141 FORMAT(/,*SIMULTANEOUS CONFIDENCE INTERVALS WILL NOW
* BE *
**PLACED ABOUT THE DIFFERENCES OF*,/,*THE MEANS*)
142 FORMAT(/,19X,*PRINCIPAL COMPONENTS WILL NOW BE
* COMPUTED*)
143 FORMAT(/,*SINCE THE TRANSFORMED DATA FOR SYSTEM *,I2,
* * HAS *
**PASSED THE TEST FOR*,/,*MULTIVARIATE NORMALITY,THIS
* TRANSFORMED *
**DATA WILL BE USED IN THE*,/,*FOLLOWING HYPOTHESIS
* TESTS*)
145 FORMAT(/,*ENTER VECTOR OF STANDARD MEANS*)
146 FORMAT(/,*VECTOR OF STANDARD MEANS IS:*,//,4(1X,F8.4))
STOP
1000 END
SUBROUTINE GEN

```

```

C
C      THIS SUBROUTINE GENERATES MVN VECTORS BASED ON
C      X=CZ+MU. COMPUTATIONS ARE ACCOMPLISHED BY USING
C      AN IMSLIB PROGRAM CALLED "GGNRM".
C

```

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
DIMENSION SIG(10),WKVEC(4),RVEC(20,4)
ISEED=486364002
DO 2 I=1,NA
M=N(I)
J=1
DO 3 L=1,IP
DO 4 K=1,L
SIG(J)=0.

```

```

      SIG(J)=SIGMA(I,L,K)
      J=J+1
4     CONTINUE
3     CONTINUE
      ISEED=ISEED+1
      IR=N(I)
      CALL GGNRM(ISEED,M,IP,SIG,IR,RVEC,WKVEC,IER)
      ISEED=466364003
      DO 6 J=1,IP
      DO 7 K=1,M
      PVAR(I,J,K)=RVEC(K,J)+SMU(I,J)
7     X(I,J,K)=RVEC(K,J)+SMU(I,J)
6     CONTINUE
2     CONTINUE
      RETURN
      END

```

SUBROUTINE MTEST

C
C
C
C

THIS SUBROUTINE TESTS A GIVEN MULTIVARIATE DATA SET
FOR MVN USING THE GENERALIZED "W-STATISTIC"

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/FOUR/US(2,20),SW(2),A(2,20)
COMMON/FIVE/YBAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SIX/AINVER(4,4),TEMPO(2,4,20),RMAX(2,20)
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
COMMON/BLOCK1/JPP,MPP,LPP,COUNT
DIMENSION SUMSQ(4,4)
INTEGER COUNT
COUNT=COUNT+1
I=IPP
M=N(I)
DO 30 J=1,IP
DO 31 K=1,M
US(I,K)=0.
RMAX(I,K)=0.
TEMPO(I,J,K)=0.
31 DIFF(I,J,K)=0.
30 CONTINUE
DO 25 J=1,IP
DO 26 K=1,IP
AINVER(J,K)=0.
26 SUMSQ(J,K)=0.

```

```

25  CONTINUE
    SUM7=0.
    DO 2 J=1,IP
    SUM6=0.
    DO 3 K=1,M
    SUM6=SUM6+X(I,J,K)
3   CONTINUE
    YBAR(I,J)=SUM6/FLOAT(M)
2   CONTINUE
    IF(COUNT.LE.2) GO TO 80
    GO TO 81
80  DO 83 LL=1,IP
83  JBAR(I,LL)=YBAR(I,LL)
81  DO 4 K=1,M
    DO 5 J=1,IP
5   DIFF(I,J,K)=X(I,J,K)-YBAR(I,J)
    DO 6 L=1,IP
    DO 7 LL=1,IP
7   SUMSQ(L,LL)=SUMSQ(L,LL)+(DIFF(I,L,K)*DIFF(I,LL,K))
6   CONTINUE
+   CONTINUE
    DO 60 J=1,IP
    DO 70 K=1,IP
    IF(COUNT.GT.2) GO TO 70
    PMAT(I,J,K)=SUMSQ(J,K)
70  AMAT(I,J,K)=SUMSQ(J,K)
60  CONTINUE
    CALL INVERS(SUMSQ,IP,IP,IPR,AINVER,D1)
    DO 8 K=1,M
    DO 9 J=1,IP
    DO 10 JJ=1,IP
10  TEMPD(I,J,K)=TEMPD(I,J,K)+(DIFF(I,JJ,K)*AINVER(JJ,J))
9   CONTINUE
8   CONTINUE
11  DO 12 K=1,M
    DO 13 J=1,IP
13  RMAX(I,K)=RMAX(I,K)+(TEMPD(I,J,K)*DIFF(I,J,K))
12  CONTINUE
    SMAX=RMAX(I,1)
    DO 14 K=2,M
    IF(RMAX(I,K).GE.SMAX) GO TO 40
    GO TO 14
+0  SMAX=RMAX(I,K)
    MAX=K
14  CONTINUE
    DO 15 J=1,M
    DO 16 K=1,IP
16  US(I,J)=US(I,J)+(TEMPD(I,K,MAX)*DIFF(I,K,J))
15  CONTINUE
    L=M-1

```

```

DO 50 J=1,L
  IP1=J+1
  DO 17 K=IP1,M
    IF(US(I,J).LE.US(I,K)) GO TO 17
    TEMP1=US(I,J)
    US(I,J)=US(I,K)
    US(I,K)=TEMP1
  17 CONTINUE
50 CONTINUE
  DO 55 J=1,M
    A(I,J)=-A(I,M-J+1)
  DO 22 J=1,M
    SUM7=SUM7+(A(I,J)*US(I,J))
  SW(I)=(SUM7**2)/RMAX(I,MAX)
22 CONTINUE
  RETURN
  END

```

SUBROUTINE TRANS(IT)

C
C
C
C
C
C
C

THIS SUBROUTINE TRANSFORMS A GIVEN MULTIVARIATE
DATA SET TO INSURE MVN AND COMPUTES THE CONFI-
DENCE LEVEL ASSOCIATED WITH THE TRANSFORMATION
PARAMETERS

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
COMMON/ELEVEN/X1(4),XSTAR(4)
COMMON/BLOCK2/YDIFF(2,4),XT(2,4),SZ1(4,4),SY1(4,4)
REAL MC
DO 250 J=1,IP
250 X1(J)=1.
  CALL FUNC(X1,F)
  G=F
  DMIN=.001
  DEL=.01
  AMDA=1.
  DO 1 I=1,IT
1  XSTAR(I)=.5
    SW=0.5
    DO 60 I=1,IT
    X1(I)=XSTAR(I)

```

```

60  CONTINUE
    ZSTAR=999999999.0
    DISP=0.
    XMOVE=3.
100  DO 170 I=1,IT
110  X1(I)=XSTAR(I)+XMOVE*DEL
    CALL FUNC(X1,F)
140  PFN=F+AMDA*FEAS
    IF(PFN.GE.ZSTAR) GO TO 150
    DISP=DISP+(ABS(XMOVE))*DEL
    XSTAR(I)=X1(I)
    XMOVE=XMOVE+XMOVE
    ZSTAR=PFN
145  FEAS=0.
    GO TO 110
150  X1(I)=XSTAR(I)
    IF(XMOVE.GT.3.3.OR.XMOVE.LT.2.7) GO TO 160
    XMOVE=-3.
    GO TO 145
160  XMOVE=3.
    FEAS=0.
170  CONTINUE
    FEAS=0.
    IF(DISP.LT.DMIN) GO TO 190
180  DISP=0.
    XMOVE=3.
    GO TO 100
190  CALL FUNC(X1,F)
    M=N(1)
    DO 200 J=1,IP
    DO 205 K=1,M
    IF(X1(J).NE.0) GO TO 210
    GO TO 215
210  X(IPP,J,K)=((X(IPP,J,K)**X1(J))-1)/X1(J)
    GO TO 205
215  X(IPP,J,K)=ALOG(X(IPP,J,K))
205  CONTINUE
200  CONTINUE
    DO 220 J=1,IP
220  XT(IPP,J)=X1(J)
    STAT=2.*(G-F)
    PARM=IP
    H=CHIFRB(STAT,PARM)
    RETURN
    END

```

SUBROUTINE FUNC(X1,F)

C
C
C
C
C

THIS SUBROUTINE EVALUATES THE LOG LIKLIHOOD
FUNCTION OF THE TRANSFORMATION PARAMETERS
COMPUTED IN SUBROUTINE TRANS

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
DIMENSION X1(4),XSTAR(4)
DIMENSION SML(4,4),SUM(4),YBAR1(4),Y(2,4,20)
DIMENSION DIFF1(4,20),SUMSQ1(4,4),IPR(4)
EXTERNAL VIPDA
M=N(1)
DO 7 J=1,IP
DO 8 K=1,M
IF(X1(J).NE.0.) GO TO 9
GO TO 10
9 Y(IPP,J,K)=((X(IPP,J,K)**X1(J))-1)/X1(J)
GO TO 8
10 Y(IPP,J,K)=ALOG(X(IPP,J,K))
8 CONTINUE
7 CONTINUE
DO 11 J=1,IP
DO 12 K=1,IP
12 SUMSQ1(J,K)=0.
11 CONTINUE
DO 13 J=1,IP
SUM6=0.
DO 14 K=1,M
14 SUM6=SUM6+Y(IPP,J,K)
YBAR1(J)=SUM6/FLOAT(M)
13 CONTINUE
DO 15 K=1,M
DO 16 J=1,IP
16 DIFF1(J,K)=Y(IPP,J,K)-YBAR1(J)
DO 17 L=1,IP
DO 18 LL=1,IP
18 SUMSQ1(L,LL)=SUMSQ1(L,LL)+(DIFF1(L,K)*DIFF1(LL,K))
17 CONTINUE
15 CONTINUE
DO 19 J=1,IP
DO 19 K=1,IP
19 SML(J,K)=SUMSQ1(J,K)/FLOAT(M)
20 CONTINUE
CALL DECOM(SML,IP,IP,IPR,IPR,D1,VIPDA)
DETSML=D1
DO 20 L=1,IP
60 DETSML=DETSML*SML(L,L)
DO 31 J=1,IP

```

```

SUM(J)=J.
DO 32 K=1,M
32 SUM(J)=SUM(J)+ALOG(PVAR(IPP,J,K))
31 CONTINUE
F1=J.
DO 33 J=1,IP
33 F1=F1+(X1(J)-1)*SUM(J)
F=(FLOAT(M)/2.)*ALOG(DETSML)-F1
RETURN
END

```

SUBROUTINE STEST

```

C
C
C   THIS SUBROUTINE TESTS EQUALITY OF TWO COVARIANCE
C   MATRICES BASED ON THE GENERALIZED BARTLETT
C   STATISTIC
C
COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/FIVE/YBAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
DIMENSION S1(4,4),IPR(4),SP1(4,4)
EXTERNAL VIPDA
REAL MC
SUML=0.
DO 5 I=1,NA
M=N(I)-1
DO 10 J=1,IP
DO 15 K=1,IP
15 S(I,J,K)=AMAT(I,J,K)/FLOAT(M)
10 CONTINUE
5 CONTINUE
DO 25 J=1,IP
DO 30 K=1,IP
SPOOL(J,K)=(S(1,J,K)+S(2,J,K))/2.
SP2(J,K)=(S(1,J,K)+S(2,J,K))/2.
30 SP1(J,K)=(S(1,J,K)+S(2,J,K))/2.
25 CONTINUE
DO 35 I=1,NA
DO 40 J=1,IP
DO 45 K=1,IP
45 S1(J,K)=S(I,J,K)
40 CONTINUE

```



```

      CALL DECOM(S1,IP,IP,IPR,IPR,D1,VIPDA)
      DETS1=D1
      DO 60 L=1,IP
60    DETS1=DETS1*S1(L,L)
      SLOG=ALOG(DETS1)
      SLOG=(N(I)-1)*SLOG
      SUML=SUML+SLOG
35    CONTINUE
      CALL DECOM(SP1,IP,IP,IPR,IPR,D1,VIPDA)
      DETSP=D1
      DO 65 L=1,IP
65    DETSP=DETSP*SP1(L,L)
      SPLOG=ALOG(DETSP)
      TEST=((2.*(N(1)-1))*SPLOG)-SUML
      INT=6*(IP+1)*NA*(N(1)-1)
      CINV=1.-(((2.*(IP**2))+(3.*IP)-1)*(NA+1))/FLOAT(INT)
      MC=TEST*CINV
      PARM=.5*IP*(IP+1)
      PROB=1-ALPHA
      F1=PICHI(PROB,PARM,IP)
      IF(MC.LE.F1) GO TO 70
      GO TO 75
70    CALL OUTPUT(7)
      NPP=1
      GO TO 100
75    CALL OUTPUT(6)
      NPP=2
100   RETURN
      END

```

SUBROUTINE MUTEST

C
C
C
C
C

THIS SUBROUTINE TESTS WHETHER A MEAN VECTOR IS
EQUAL TO A STANDARD MEAN VECTOR AND WHETHER THE
DIFFERENCE BETWEEN TWO MEAN VECTORS IS ZERO

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISO,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/FIVE/YBAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
COMMON/BLOCK1/JPP,MPP,LPP,COUNT
COMMON/BLOCK2/YDIFF(2,4),XT(2,4),SZ1(4,4),SY1(4,4)
DIMENSION SPINV(4,4),TEMPS(4),DIFFS(4),DF(2),IPR(4)

```

```

    DIMENSION BFINV(4,4),SUMSQ3(4,4),TBF(4),Z(4,20)
    DIMENSION BFDIFF(4,20),CFS(4,4)
    DIMENSION SY(4,4),SZ(4,4),SYINV(4,4),SZINV(4,4),
*   YTEMP(4)
    REAL MC
    JPP=0
    MPD=0
    IF(IK.EQ.1) GO TO 299
    DO 200 I=1,NA
    DO 205 J=1,IP
205  YDIFF(I,J)=YBAR(I,J)-(((XMU(J)**XT(I,J))-1)/XT(I,J))
200  CONTINUE
    GO TO 350
299  DO 300 I=1,NA
    DO 305 J=1,IP
305  YDIFF(I,J)=YBAR(I,J)-XMU(J)
300  CONTINUE
350  DO 215 J=1,IP
    DO 220 K=1,IP
    SZ(J,K)=S(1,J,K)
    SZ1(J,K)=S(1,J,K)
    SY(J,K)=S(2,J,K)
220  SY1(J,K)=S(2,J,K)
215  CONTINUE
    DO 210 I=1,NA
    IF(I.EQ.1) GO TO 216
    GO TO 217
216  CALL INVERS(SZ,IP,IP,IPR,SZINV,D1)
    GO TO 218
217  CALL INVERS(SY,IP,IP,IPR,SYINV,D1)
218  DO 225 J=1,IP
    YTEMP(J)=0.
    DO 230 K=1,IP
    IF(I.EQ.1) GO TO 226
    GO TO 227
226  YTEMP(J)=YTEMP(J)+(YDIFF(I,K)*SZINV(K,J))
    GO TO 230
227  YTEMP(J)=YTEMP(J)+(YDIFF(I,K)*SYINV(K,J))
230  CONTINUE
225  CONTINUE
    YMAT=0.
    DO 235 J=1,IP
    YMAT=YMAT+(YTEMP(J)*YDIFF(I,J))
235  CONTINUE
    TSQ=N(I)*YMAT
    DF(1)=IP
    DF(2)=N(I)-IP
    PROB=1.-ALPHA
    F=PIFDIS(PROB,DF,IR)
    YPROB1=FLOAT(IP*(N(I)-1))

```

```

YPRD2=FLOAT(N(I)-IP)
TNOTSQ=(YPRD1/YPRD2)*F
IF(TSQ.LE.TNOTSQ) GO TO 240
GO TO 245
240 LPP=I
CALL OUTPUT(19)
GO TO 210
245 LPP=I
CALL OUTPUT(20)
IF(I.EQ.1) JPP=1
IF(I.EQ.2) MPP=1
210 CONTINUE
IF(NPP.EQ.1) GO TO 1
IF(NPP.EQ.2) GO TO 50
1 SMAT=0.
DO 2 J=1,IP
2 TEMPS(J)=0.
CALL INVERS(SP2,IP,IP,IPR,SPINV,D1)
DO 10 K=1,IP
10 DIFFS(K)=YBAR(1,K)-YBAR(2,K)
DO 15 J=1,IP
DO 20 K=1,IP
20 TEMPS(J)=TEMPS(J)+(DIFFS(K)*SPINV(K,J))
15 CONTINUE
DO 30 K=1,IP
30 SMAT=SMAT+(TEMPS(K)*DIFFS(K))
TSQ=((N(1)*N(2))/FLOAT(N(1)+N(2)))*SMAT
DF(1)=IP
DF(2)=N(1)+N(2)-IP-1
PROB=1-ALPHA
F2=PIFDIS(PROB,DF,IR)
PROD1=FLOAT(N(1)+N(2)-2)
PROD2=FLOAT(N(1)+N(2)-IP-1)
TNOTSQ=(PROD1/PROD2)*F2
32 IF(TSQ.LE.TNOTSQ) GO TO 35
GO TO 40
35 CALL OUTPUT(9)
KPP=1
GO TO 100
40 CALL OUTPUT(8)
KPP=2
GO TO 100
50 DO 55 J=1,IP
DO 56 K=1,IP
BFINV(J,K)=0.
SUMSQ3(J,K)=0.
56 TBF(J)=0.
55 CONTINUE
DO 60 J=1,IP
M=N(1)

```

```

      DO 65 K=1,M
65   Z(J,K)=X(1,J,K)-X(2,J,K)
60   CONTINUE
      DO 70 J=1,IP
      M=N(1)
      SUM10=0.
      DO 75 K=1,M
      SUM10=SUM10+Z(J,K)
75   CONTINUE
      ZBAR(J)=SUM10/FLOAT(M)
70   CONTINUE
      DO 80 K=1,M
      DO 85 J=1,IP
85   BFDIFF(J,K)=Z(J,K)-ZBAR(J)
      DO 90 L=1,IP
      DO 95 LL=1,IP
85   SUMSQ3(L,LL)=SUMSQ3(L,LL)+(BFDIFF(L,K)*BFDIFF(LL,K))
90   CONTINUE
80   CONTINUE
      M=N(1)-1
      DO 105 J=1,IP
      DO 110 K=1,IP
      CFS(J,K)=SUMSQ3(J,K)/FLOAT(M)
110  BFS(J,K)=SUMSQ3(J,K)/FLOAT(M)
105  CONTINUE
      CALL INVERS(CFS,IP,IP,IPR,BFINV,D1)
      DO 115 J=1,IP
      DO 120 K=1,IP
120  TBF(J)=TBF(J)+(ZBAR(K)*BFINV(K,J))
115  CONTINUE
      DFS=0.
      DO 125 J=1,IP
125  DFS=DFS+(TBF(J)*ZBAR(J))
      TSQ=N(1)*DFS
      DF(1)=IP
      DF(2)=N(1)-IP
      PROB=1-ALPHA
      F2=PIFDIS(PROB,DF,IR)
      PROD1=FLOAT((N(1)-1)*IP)
      PROD2=FLOAT(N(1)-IP)
      TNOTSQ=(PROD1/PROD2)*F2
      GO TO 32
100  RETURN
      END

```

SUBROUTINE CONINT

THIS SUBROUTINE CALCULATES SIMULTANEOUS CONFIDENCE
INTERVALS ABOUT MEANS AND DIFFERENCES OF MEANS IF
DETECTED BY SUBROUTINE MUTEST

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISC,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/FIVE/YBAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
COMMON/TEN/CLIML(4),CLIMU(4)
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
COMMON/BLOCK1/JPP,MPP,LPP,COUNT
COMMON/BLOCK2/YDIFF(2,4),XT(2,4),SZ1(4,4),SY1(4,4)
DIMENSION DEV(4),DF(2),XDEV(4),YDEV(4),XD(4),YD(4),
*  QL(4),QR(4)
REAL MC
IF(JPP.EQ.0.AND.MPP.EQ.0) GO TO 5
DF(1)=FLOAT(IP)
DF(2)=FLOAT(N(1)-IP)
PROB=1.-ALPHA
F=PIFDIS(PROB,DF,IR)
Q1=FLOAT(N(1)-1)/FLOAT(N(1))
CONST=SQRT(Q1*(DF(1)/DF(2))*F)
IF(JPP.EQ.1) GO TO 40
IF(MPP.EQ.1) GO TO 45
40 DO 50 J=1,IP
XD(J)=SQRT(PMAT(1,J,J)/(N(1)-1))
50 XDEV(J)=SQRT(SZ1(J,J))
DO 55 KK=1,IP
QL(KK)=QBAR(1,KK)-(XD(KK)*CONST)
QR(KK)=QBAR(1,KK)+(XD(KK)*CONST)
CLIML(KK)=YDIFF(1,KK)-(XDEV(KK)*CONST)
CLIMU(KK)=YDIFF(1,KK)+(XDEV(KK)*CONST)
55 CONTINUE
LPP=1
CALL OUTPUT(10)
+5 IF(MPP.EQ.0) GO TO 5
DO 60 J=1,IP
YD(J)=SQRT(PMAT(2,J,J)/(N(2)-1))
60 YDEV(J)=SQRT(SY1(J,J))
DO 65 KK=1,IP
QL(KK)=QBAR(2,KK)-(YD(KK)*CONST)
QR(KK)=QBAR(2,KK)+(YD(KK)*CONST)
CLIML(KK)=YDIFF(2,KK)-(YDEV(KK)*CONST)
CLIMU(KK)=YDIFF(2,KK)+(YDEV(KK)*CONST)
65 CONTINUE
LPP=2
CALL OUTPUT(10)

```

```

5  IF(KPP.EQ.2.AND.NPP.EQ.2) GO TO 20
   IF(KPP.EQ.1.AND.NPP.EQ.1) GO TO 100
   IF(KPP.EQ.1.AND.NPP.EQ.2) GO TO 100
   PROD1=(N(1)+N(2))*(N(1)+N(2)-2)*IP
   PROD2=(N(1)*N(2))*(N(1)+N(2)-IP-1)
   CONST=SQRT((PROD1/PROD2)*F2)
   DO 10 J=1,IP
10  DEV(J)=SQRT(SPOOL(J,J))
   DO 15 KK=1,IP
   CLIML(KK)=YBAR(1,KK)-YBAR(2,KK)-(DEV(KK)*CONST)
   CLIMU(KK)=YBAR(1,KK)-YBAR(2,KK)+(DEV(KK)*CONST)
15  CONTINUE
   CALL OUTPUT(10)
   GO TO 100
20  DO 30 J=1,IP
30  DEV(J)=SQRT(BFS(J,J))
   DO 35 KK=1,IP
   CLIML(KK)=ZBAR(KK)-(DEV(KK)*CONST)
   CLIMU(KK)=ZBAR(KK)+(DEV(KK)*CONST)
35  CONTINUE
   LPP=3
   CALL OUTPUT(10)
100 RETURN
   END

```

SUBROUTINE PRICOM

C
C
C
C
C
C
C
C
C
C

THIS SUBROUTINE EXTRACTS PRINCIPAL COMPONENTS FROM
THE COVARIANCE MATRIX OR CORRELATION MATRIX OF A
GIVEN DATA SET AND THE DIFFERENCE OF TWO DATA SETS.
IT ALSO CALCULATES THE CONTRIBUTION OF INDIVIDUAL
MOE TO THE OPERATIONAL EFFECTIVENESS OF A FORCE

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISO,IK,KI,SUM5,I3,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/NINE/EVAL(4),EVEC(4,4),VAR(4),XCONT(4)
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
DIMENSION COR(4,4),Y1(4,4),Y2(4,20),Y2BAR(4),COV(4,4),
* YY(10)
DIMENSION SUMSQ2(4,4),TODIFF(4,20),RC(10),CC(10)
DIMENSION COMP(4,4),CL(4),CU(4),VECSQ(4,4)

```

```

DO 30 I=1,NA
M=N(I)
MI=I
CL(1)=20
DO 35 J=1,IP
DO 40 K=1,IP
40 Y1(J,K)=PMAT(I,J,K)/FLOAT(M-1)
35 CONTINUE
DO 65 J=1,IP
DO 70 K=1,IP
70 COR(J,K)=(Y1(J,K))/((SQRT(Y1(J,J)))*(SQRT(Y1(K,K))))
65 CONTINUE
L=1
DO 135 J=1,IP
DO 140 K=1,J
YY(L)=0.
RC(L)=0.
YY(L)=Y1(J,K)
RC(L)=COR(J,K)
L=L+1
140 CONTINUE
135 CONTINUE
IF (IA.EQ.1) GO TO 5
GO TO 10
5 CALL OPRINC(YY,IP,IP,EVAL,EVEC,COMP,VAR,CL,CU,IER)
GO TO 15
10 CALL OPRINC(RC,IP,IP,EVAL,EVEC,COMP,VAR,CL,CU,IER)
15 L=IP
DO 250 J=1,IP
DO 260 K=1,IP
260 VECSQ(J,K)=VAR(L)*(EVEC(K,J)**2)
L=IP-J
250 CONTINUE
DO 270 J=1,IP
XCONT(J)=0.
DO 280 K=1,IP
280 XCONT(J)=XCONT(J)+VECSQ(K,J)
270 CONTINUE
CALL OUTPUT(11)
30 CONTINUE
M=N(1)
DO 45 J=1,IP
DO 50 K=1,M
50 Y2(J,K)=PVAR(1,J,K)-PVAR(2,J,K)
45 CONTINUE
DO 46 J=1,IP
DO 47 K=1,IP
47 SUMSQ2(J,K)=0.
46 CONTINUE
DO 75 J=1,IP

```

```

      M=N(1)
      SUM9=0.
      DO 80 K=1,M
      SUM9=SUM9+Y2(J,K)
80    CONTINUE
      Y2BAR(J)=SUM9/FLOAT(M)
75    CONTINUE
      DO 85 K=1,M
      DO 90 J=1,IP
80    TDIFF(J,K)=Y2(J,K)-Y2BAR(J)
      DO 95 L=1,IP
      DO 100 LL=1,IP
100   SUMSQ2(L,LL)=SUMSQ2(L,LL)+(TDIFF(L,K)*TDIFF(LL,K))
95    CONTINUE
85    CONTINUE
      M=N(1)-1
      DO 105 J=1,IP
      DO 110 K=1,IP
110   COV(J,K)=SUMSQ2(J,K)/FLOAT(M)
105   CONTINUE
      L=1
      DO 235 J=1,IP
      DO 240 K=1,J
      CC(L)=0.
      CC(L)=COV(J,K)
      L=L+1
240   CONTINUE
235   CONTINUE
      CL(1)=20
      CALL CPRINC(CC,IP,IP,EVAL,EVEC,COMP,VAR,CL,CU,IER)
      L=IP
      DO 350 J=1,IP
      DO 360 K=1,IP
360   VECSQ(J,K)=VAR(L)*(EVEC(K,J)**2)
      L=IP-J
350   CONTINUE
      DO 370 J=1,IP
      XCONT(J)=0.
      DO 380 K=1,IP
380   XCONT(J)=XCONT(J)+VECSQ(K,J)
370   CONTINUE
      CALL OUTPUT(18)
      RETURN
      END

```


SUBROUTINE OUTPUT(K)

C
C
C
C

THIS SUBROUTINE PRINTS OUTPUT COMPUTED IN OTHER
SUBROUTINES

```

COMMON/ONE/X(2,4,20),SIGMA(2,4,4),N(2),SMU(2,4)
COMMON/TWO/IP,NA,CHISQ,IK,KI,SUM5,IB,IPP,KPP,NPP,IA,
COMMON/THREE/PVAR(2,4,20)
COMMON/FOUR/US(2,20),SW(2),A(2,20)
COMMON/FIVE/YPAR(2,4),DIFF(2,4,20),IPR(20),AMAT(2,4,4)
COMMON/SIX/AINVER(4,4),TEMPD(2,4,20),RMAX(2,20)
COMMON/SEVEN/S(2,4,4),SPOOL(4,4),BFS(4,4),SP2(4,4)
COMMON/EIGHT/ALPHA,F1,F2,MC,TSQ,TNOTSQ,H,STAT
COMMON/NINE/EVAL(4),EVEC(4,4),VAR(4),XCONT(4)
COMMON/TEN/CLIML(4),CLIMU(4)
COMMON/ELEVEN/X1(4),XSTAR(4)
COMMON/TWELVE/XMU(4),PMAT(2,4,4),ZBAR(4),QBAR(2,4)
COMMON/BLOCK1/JPP,MPP,LPP,COUNT
COMMON/BLOCK2/YDIFF(2,4),XT(2,4),SZ1(4,4),SY1(4,4)
REAL MC
IF(K.EQ.1) GO TO 1
IF(K.EQ.2) GO TO 2
IF(K.EQ.3) GO TO 3
IF(K.EQ.4) GO TO 4
IF(K.EQ.5) GO TO 5
IF(K.EQ.6) GO TO 6
IF(K.EQ.7) GO TO 7
IF(K.EQ.8) GO TO 8
IF(K.EQ.9) GO TO 9
IF(K.EQ.10) GO TO 10
IF(K.EQ.11) GO TO 11
IF(K.EQ.18) GO TO 18
IF(K.EQ.19) GO TO 19
IF(K.EQ.20) GO TO 20
1 CONTINUE
2 CONTINUE
3 CONTINUE
4 IF(IPP.EQ.1) WRITE(6,27) (IPP, SW(IPP))
  IF(IPP.EQ.2) WRITE(6,27) (IPP, SW(IPP))
  IF(IPP.EQ.0) WRITE(6,27) (I, SW(I),I=1,NA)
  GO TO 100
5 WRITE(6,45) IPP
  WRITE(6,28) (I,X1(I),I=1,IP)
  WRITE(6,46) STAT,H
  GO TO 100
6 WRITE(6,29) MC,F1
  WRITE(6,30)
  GO TO 100
7 WRITE(6,29) MC,F1
  WRITE(6,31)

```

```

      GO TO 100
8    WRITE(6,48)
      WRITE(6,32)  TSQ,TNOTSQ
      WRITE(6,30)
      GO TO 100
9    WRITE(6,48)
      WRITE(6,32)  TSQ,TNOTSQ
      WRITE(6,31)
      GO TO 100
10   IF(LPP.EQ.3)  GO TO 90
      WRITE(6,33)  LPP
      WRITE(6,34)  (CLIML(L),CLIMU(L),L=1,IP)
      GO TO 100
90   WRITE(6,53)
      WRITE(6,34)  (CLIML(L),CLIMU(L),L=1,IP)
      GO TO 100
11   WRITE(6,35)  MI, (EVAL(J),J=1,IP)
      WRITE(6,36)  MI
      WRITE(6,37)  ((EVEC(J,L),L=1,IP),J=1,IP)
      WRITE(6,38)  MI, (VAR(J),J=1,IP)
      WRITE(6,49)  MI
      WRITE(6,50)
      WRITE(6,51)  (I,XCONT(I),I=1,IP)
      WRITE(6,44)
      GO TO 100
18   WRITE(6,40)  (EVAL(J),J=1,IP)
      WRITE(6,41)
      WRITE(6,42)  ((EVEC(J,L),L=1,IP),J=1,IP)
      WRITE(6,43)  (VAR(J),J=1,IP)
      WRITE(6,52)
      WRITE(6,50)
      WRITE(6,51)  (I,XCONT(I),I=1,IP)
      GO TO 100
19   WRITE(6,47)  LPP
      WRITE(6,32)  TSQ,TNOTSQ
      WRITE(6,31)
      GO TO 100
20   WRITE(6,47)  LPP
      WRITE(6,32)  TSQ,TNOTSQ
      WRITE(6,30)
21   FORMAT(F15.9)
22   FORMAT(/,3X,*CHI-SQUARED STATISTIC = *,F7.4)
23   FORMAT(/,3X,*IK = *,I3)
24   FORMAT(/,*C-MATRIX FOR SYSTEM *,I2,* IS:*,/)
25   FORMAT(3X,8(1X,F8.4))
26   FORMAT(/,*OBSERVATIONS FOR SYSTEM *,I2,* ARE:*,/)
27   FORMAT(/,*THE SHAPIRO-WILK STATISTIC FOR SYSTEM*,I2,*
*   IS: *,F7.4)
28   FORMAT(+ (28X,*LAMBDA(*,I2,*) = *,F7.4,/))
29   FORMAT(/,*BARTLETT'S STATISTIC = *,F9.4,5X,*CHI-SQ.
*   STATISTIC = *,

```

```

*F7.4)
30  FORMAT(/,*THE NULL HYPOTHESIS IS REJECTED*)
31  FORMAT(/,*WE FAIL TO REJECT THE NULL HYPOTHESIS*)
32  FORMAT(/,*T-SQ. STATISTIC = *,F9.4,5X,*HOTELLING TEST
*   STATISTIC
*   = *,F7.4)
33  FORMAT(/,17X,*SYSTEM *,I2,* MEAN VECTOR AND STANDARD
**MEAN VECTOR*,/,25X,*LOWER LIMIT*,9X,*UPPER LIMIT*)
34  FORMAT(/,4(27X,F7.4,13X,F7.4,/) )
35  FORMAT(/,21X,*EIGENVALUES FOR SYSTEM *,I2,* DATA
*   MATRIX*,//,
*20X,4(1X,F7.4))
36  FORMAT(/,20X,*EIGENVECTORS FOR SYSTEM *,I1,* DATA
*   MATRIX*,/)
37  FORMAT(4(21X,4(1X,F7.4),/))
38  FORMAT(/,5X,*PERCENT OF SYSTEM *,I2,* VARIANCE
*   ATTRIBUTED *
**TO ITS PRINCIPAL COMPONENTS*,//,20X,4(1X,F7.4),/)
40  FORMAT(/,20X,*EIGENVALUES FOR DIFFERENCE MATRIX*,//,
*20X,4(1X,F7.4))
44  FORMAT(/,10X,58(1H*),/)
41  FORMAT(/,20X,*EIGENVECTORS FOR DIFFERENCE MATRIX*,//)
42  FORMAT(4(21X,4(1X,F7.4),/))
43  FORMAT(/,*PERCENT VARIANCE OF DIFFERENCE MATRIX *
**ATTRIBUTED TO ITS PRINCIPAL COMPONENTS*,//,20X,4(1X,
*   F7.4))
45  FORMAT(/,20X,*TRANSFORMATION PARAMETERS FOR SYSTEM *,
*   I2,/)
46  FORMAT(/,*LIKLIHOOD RATIO STATISTIC = *,F7.4,3X,*CON*
**FIDENCE LEVEL = *,F6.4)
47  FORMAT(/,*TEST FOR EQUALITY OF SYSTEM *,I2,* MEAN
*   VECTOR *
**AND STANDARD MEAN VECTOR*)
48  FORMAT(/,*TEST FOR EQUALITY OF SYSTEM 1 AND SYSTEM 2
*   MEAN *
**VECTORS*)
49  FORMAT(/,*PERCENT CONTRIBUTION OF INDIVIDUAL MOE TO *
**OPERATIONAL EFFECTIVENESS*,/,*WITH SYSTEM *,I2)
50  FORMAT(/,22X,*MOE*,15X,*CONTRIBUTION*)
51  FORMAT(/,4(22X,I2,19X,F7.4,/) )
52  FORMAT(/,*PERCENT CONTRIBUTION OF DIFFERENCES IN MOE
**TO DIFFERENCES IN OPERATIONAL*,/,*EFFECTIVENESS*)
53  FORMAT(/,23X,*SYSTEM 1 AND SYSTEM 2 MEAN VECTORS *
*,/,25X,*LOWER LIMIT*,9X,*UPPER LIMIT*)
100 RETURN
END

```

BIBLIOGRAPHY

1. Alt, F. B., "Aspects of Multivariate Control Charts", unpublished Master's thesis, Georgia Institute of Technology, 1973.
2. Andrews, M., H. A. Crosby and T. N. DuPuy, "Developing a Methodology to Describe the Relationship of Mobility to Combat Effectiveness", Research Analysis Corporation, Washington, D. C., 1967.
3. Andrews, D. F., R. Gnanadesikan and J. L. Warner, "Transformations of Multivariate Data", Biometrics, Vol. 27, 1971, pp. 825-840.
4. Andrews, D. F., "A Note on the Selection of Data Transformations", Biometrika, Vol. 58, No. 2, pp. 249-254.
5. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York, 1958.
6. Bazaraa, M. S., An Efficient Cyclic Coordinate Method for Optimizing Penalty Functions, unpublished report, Georgia Institute of Technology, 1975.
7. Bennett, B. M., "Note on a Solution of the Generalized Behrens-Fisher Problem", Annals of Mathematical Statistics, Vol. 2, 1951, pp. 87-90.
8. Bennett, M., "On a Certain Multivariate Non-Normal Distribution", Proceedings of the Cambridge Philosophical Society, Vol. 57, 1961, pp. 434-436.
9. Bettencourt, V. M., "An Application of Multiple Response Surface Optimization to the Analysis of Training Effects in Operational Testing", unpublished Master's thesis, Georgia Institute of Technology, 1975.
10. Birren, J. E. and D. F. Morrison, "Analysis of the WAIS Subtests in Relation to Age and Education", Journal of Gerontology, Vol. 16, 1961, pp. 363-369.
11. Box, G. E. P., "A General Distribution Theory for a Glass of Likelihood Criteria", Biometrika, Vol. 36, 1949, pp. 317-346.
12. Box, G. E. P. and D. R. Cox, "An Analysis of Transformations", Journal of the Royal Statistical Society of Britain, Series B, Vol. 26, 1964, pp. 211-252.

13. Burnette, T. H., Jr., "A Comparison of the Applicability and Effectiveness of ANOVA and MANOVA for Use in the Operational Evaluation of Command and Control Systems", unpublished Master's thesis, Georgia Institute of Technology, 1975.
14. Fishman, G. S., Concepts and Methods in Discrete Event Digital Simulation, John Wiley and Sons, New York, 1973
15. Gartenhaus, S., Physics: Basic Principles, Vol. 1, Holt, Rinehart, and Winston, Inc., New York, 1975.
16. Girshick, M. A., "Principal Components", Journal of the American Statistical Association, Vol. 31, 1936, pp. 519-528.
17. Graybill, F. A., An Introduction to Linear Statistical Models, Vol. 1, McGraw-Hill Book Company, New York, 1961.
18. Hines, W. W. and D. C. Montgomery, Probability and Statistics in Engineering and Management Science, The Ronald Press, New York, 1972.
19. Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components", Journal of Educational Psychology, Vol. 24, 1933, pp. 417-441, 498-520.
20. Hotelling, H., "Simplified Calculation of Principal Components", Psychometrika, Vol. 1, 1936, pp. 27-35.
21. Hotelling, H., "The Generalization of Student's Ratio", Annals of Mathematical Statistics, Vol. 2, 1931, pp. 360-378.
22. John, S., "Some Optimal Multivariate Tests", Biometrika, Vol. 58, No. 1, 1971, pp. 123-127.
23. Jolicoeur, P. and J. E. Mosimann, "Size and Shape Variation in the Painted Turtle: A Principal Component Analysis", Growth, Vol. 24, 1960, pp. 339-354.
24. Jurkat, M. P., C. J. Nuttall and P. W. Haley, "The AMC '74 Mobility Model", Technical Report No. 11921(LL-149), U. S. Army Tank Automotive Command, Mobility Systems Laboratory, Warren, Michigan, May 1975.
25. Kendall, M. G., A Course in Multivariate Analysis, Charles Griffin Company, Ltd., London, 1963.
26. Lockhart, R. S., "The Assumption of Multivariate Normality", British Journal of Mathematical and Statistical Psychology, Vol. 20, 1967, pp. 63-69.

27. Malkovich, J. F. and A. A. Afifi, "On Tests for Multivariate Normality", Journal of the American Statistical Association, Vol. 68, No. 341, 1973, pp. 176-179.
28. Mardia, K. V., "Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies", Sankhya: The Indian Journal of Statistics, Series B, Vol. 36, Part 2, 1974, pp. 115-128.
29. Mardia, K. V., "Measures of Multivariate Skewness and Kurtosis with Applications", Biometrika, Vol. 57, 1970, pp. 519-530.
30. Morrison, D. F., Multivariate Statistical Methods, McGraw-Hill Book Company, New York, 1967.
31. Pearson, K., "On Lines and Planes of Closest Fit to Systems of Points in Space", Philosophical Magazine, Serial 6, Vol. 2, 1901, pp. 559-572.
32. Press, S. James, Applied Multivariate Analysis, Holt, Rinehart, and Winston, Inc., New York, 1972.
33. Rao, C. R., Linear Statistical Inference and Its Applications, 2nd edition, John Wiley and Sons, New York, 1973.
34. Rao, C. R., "The Use and Interpretation of Principal Component Analysis in Applied Research", Sankhya: The Indian Journal of Statistics, Series A, Vol. 26, 1964, pp. 329-358.
35. Reklaitas, G. V. and D. T. Philips, "A Survey of Nonlinear Programming", American Institute of Industrial Engineering Transactions, Vol. 7, No. 3, September, 1975, pp. 235-256.
36. Sarhan, A. E. and B. G. Greenberg, "Estimation of Location and Scale Parameters by Order Statistics from Singly and Double Censored Samples, Part I", Annals of Mathematical Statistics, Vol. 27, 1956, pp. 427-451.
37. Scheur, E. and D. S. Stoller, "On the Generation of Normal Random Vectors", Technometrics, Vol. 4, 1962, pp. 278-281.
38. Shapiro, S. S. and M. B. Wilk, "An Analysis of Variance Test for Normality (Complete Samples)", Biometrika, Vol. 52, 1965, pp. 591-611.
39. Shapiro, S. S., M. B. Wilk and H. J. Chen, "A Comparative Study of Various Tests for Normality", Journal of the American Statistical Association, Vol. 63, 1968, pp. 1343-1372.
40. Sheffé, H., "On Solutions of the Behrens-Fisher Problem Based on the t-distribution", Annals of Mathematical Statistics, Vol. 14, 1943, pp. 35-44.

UNSIGNED REFERENCES

41. AR 1000.1, dated 5 November 1974, Basic Policies for Systems Acquisition by the Department of the Army, Headquarters, Department of the Army, Washington, D. C.
42. "Combat Development Study Plan: Mechanized Infantry Combat Vehicle (MICV) Cost and Operational Effectiveness Analysis (COEA)", dated 14 August 1975, Headquarters, U. S. Army Infantry School, Fort Benning, Georgia.
43. "Computer Subroutine Libraries in Mathematics and Statistics", International Mathematical and Statistical Libraries, Inc., Houston, Texas, 1975.
44. DoD Directive 5000.1, dated 1 September 1974, Acquisition of Major Defense Systems, Headquarters, Department of Defense, Washington, D. C.
45. FM 100-5, dated September 1968, Operations of Army Forces in the Field, Headquarters, Department of the Army, Washington, D. C.
46. "Math Science Library", Control Data Corporation, Sunnyvale, Calif., 1973 (revised).
47. "Proceedings of the Twentieth Conference on the Design of Experiments in Army Research Development and Testing", U. S. Army Research Office, Report No. 75-2, 11 June 1975, Research Triangle Park, North Carolina.
48. TRADOC Regulation 11-8, dated 28 April 1975, Cost and Operational Effectiveness Analysis in the Materiel Acquisition Process, Headquarters, U. S. Army Training and Doctrine Command, Fort Monroe, Virginia.